

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07H 21/04, C07K 1/00, C12N 1/20, C12P 21/06</b>		<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/11014</b> <b>(43) International Publication Date:</b> 2 March 2000 (02.03.00)
<b>(21) International Application Number:</b> PCT/US99/19330 <b>(22) International Filing Date:</b> 24 August 1999 (24.08.99)  <b>(30) Priority Data:</b> 60/097,917 25 August 1998 (25.08.98) US 60/098,634 31 August 1998 (31.08.98) US  <b>(71) Applicant (for all designated States except US):</b> HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). SHI, Yang-gu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, MD 22020 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). EN-			<b>DRESS, Gregory, A.</b> [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). <b>EBNER, Reinhard</b> [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). <b>KOMATSOULIS, George</b> [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US). <b>DUAN, Roxanne, D.</b> [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US).  <b>(74) Agents:</b> HOOVER, Kenley, K. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).  <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> 49 HUMAN SECRETED PROTEINS			
<b>(57) Abstract</b>  The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## 49 Human Secreted Proteins

### *Field of the Invention*

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

### *Background of the Invention*

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of

the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

5

### ***Summary of the Invention***

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant and synthetic methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders and conditions related to the polypeptides and polynucleotides, and therapeutic methods for treating such disorders and conditions. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

### ***Detailed Description***

#### **Definitions**

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide/sequences of the present invention.

30



In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released  
5 into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000  
10 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding  
15 sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid  
20 sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants  
25 of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig  
30 analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the

ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

5           A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750  
10   mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

          Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions.  
15   Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M  
20   NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

          Note that variations in the above conditions may be accomplished through the  
25   inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above,  
30   due to problems with compatibility.

          Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using dT as a primer).

5           The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of  
10 single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA  
15 backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

          The polypeptide of the present invention can be composed of amino acids  
20 joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more  
25 detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of  
30 modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or

may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

**Polynucleotides and Polypeptides of the Invention****FEATURES OF PROTEIN ENCODED BY GENE NO: 1**

The translation product of this gene shares sequence homology with the TAP2 of *Rattus norvegicus* (See, e.g., Genbank Accession No. gi|407479) AAD15830; all  
 5 references available through this accession are hereby incorporated by reference herein.) which is thought to be important in antigen presentation in T-cells. As such, the protein product of this gene is useful for a variety of diagnostic tests for various immune system dysfunctions, and for intervention of the antigen presentation process (e.g., to enhance the immune response to vaccines and diminish the immune response  
 10 associated with autoimmune disease).

The major histocompatibility complex (Mhc) regions of mice, rats, and humans all contain a pair of related genes, TAP1 and TAP2, which encode members of a large superfamily of proteins of similar structure and function. A functional TAP1/TAP2 heterodimer is probably required for efficient presentation of antigens to  
 15 CD8(+) T cells. This heterodimer resides in the membrane of the endoplasmic reticulum, and transports peptides from the cytoplasm into the endoplasmic reticulum lumen for binding to Mhc class I molecules.

The translation product of this gene also shares sequence homology with the transport-associated proteins and ATP-binding proteins (see, e.g., Genbank accession  
 20 numbers CAB05918 (z83328.1) and AAB95060 (AF040659.1); all references available through these accessions are hereby incorporated by reference herein.)

Preferred polypeptides of the invention comprise the following amino acid sequence:

EPHRGPHLPPDLGHHHGQRPGQLQNINFLRNTVKVTGVVVFMSLSWQLSLV  
 25 TFMGFPIIMMVSNIYGKYYKRLSKEVQNALARASNTAEETISAMKTVRSFAN  
 EEEEEAEVYLRKLQQVYKLNKEAAA YMYVWGSGLTLLVVQVSILYYGGH  
 LVISGQMTSGNLI AFIIYEFVLGDCMENV SFSLSPGKVTALVGPSGSGKSSCVN  
 ILENFYPLEGGRVLLDGKPISAYDHKYLHRVISLVSQEPVLFARSITDNISYGLP  
 TVPFEMVVEAAQKANAHGFIMELQDGYSTETGEKGAQLSGGQKQRVAWPG  
 30 LWCGTPQSSSWMKPPALWMPRASI (SEQ ID NO: 123);  
 MSSATWTAASWRTSATSTSLTRCWISGQPACTAAACCWGATIGVAKNSALG  
 PRRLRASWL VITLVCLFVGIYAMVKLLLFSEVRRPIRDPFWFVWVWYISLG

ASFLWLLSTVRPGTQALEPGAATEAEGFPGSGRPPPEQASGATLQKLLSYT  
KPDVAFLVAASFFLIVAALGETFLPYTGRAIDGIVIQKSMDQFSTAVVIVCLL  
AIGSSFAAGIRGGIFTLIFARLNIRLNCLFRSLVSQETSFFDENRTGDLISRLTS  
DTTMVSDLVSRTSMSSCGTQSRRAWSSCSASHGSSPWSPSWASPSS (SEQ

5 ID NO: 124) or

HLLRPAHCAFRDGGGGRTEGQCPRLLHHGTPGRLQHRDRGEGRPVVRWPEAA  
GGMARALVRNPPVLILDEATSALDAESEYLIQQAIGHNLQKHTVLIHAHRLST  
VEHAHLIVVLDKGRVVQQGTHQQLLAQGGLYAKLVQRQMLGLQPAADFTA  
GHNEPVANGSHKA (SEQ ID NO: 125). Polynucleotides encoding these

10 polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

15 This gene is expressed primarily in brain and testes, and to a lesser extent in amniotic cells, merkel cells and fetal tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, reproductive, or immune disorders, particularly

20 immunodeficiency, infection, lymphomas, auto-immunities, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in

25 certain tissues or cell types (e.g., immune, hematopoietic, neural, reproductive, developmental, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or

30 bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 67 as residues: His-65 to Cys-73, His-144 to Gly-152. Polynucleotides encoding said polypeptides are also provided.

The major histocompatibility complex (Mhc) regions of mice, rats, and  
5 humans all contain a pair of related genes, TAP1 and TAP2, which encode members of a large superfamily of proteins of similar structure and function. A functional TAP1/TAP2 heterodimer is probably required for efficient presentation of antigens to CD8(+) T cells. Furthermore, the tissue distribution in merkel cells and homology to TAP indicates that polynucleotides and polypeptides corresponding to this gene are  
10 useful for the diagnosis and treatment of conditions in which antigen presentation is at issue, such as general microbial infection, auto-immunity, inflammation or cancer. Alternatively, expression within brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions  
15 such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses ,  
20 autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Furthermore, the protein may also be used to determine biological activity, to raise  
25 antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly  
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3275 of SEQ ID NO:11, b is an integer of 15 to 3289, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this gene shares sequence homology with a human retrovirus related env polypeptide pseudogene. This similarity indicates that the human protein described herein is also sent to the cell surface.

Preferred polypeptides of the invention comprise the following amino acid sequence:

RLTKTISFSLQNQTAFINSLAKTPYQALTGAALAGSYPIWENENTLSWYLPSPPTLLSPPVLCVQLIFXLPANWSGTCTLVFQAPTINILPPNQITILISVEASISSSPIR  
NKWALHLITLLTGLGITAALGTGIAGITTSITSYQTLFTTSLNTVEDMHTSITSL  
QR

QLDFLVGVILQNWRVLDLLTTEKGGTCIYLQECCFCVNESGIVHIAVRRRLHDDRAAEL (SEQ ID NO: 126), YPIWENENTLSWYLPSPPTLLSPPVLCV (SEQ ID NO: 127), or RVLDDLTTTEKGGTCIYLQECCFCVNE (SEQ ID NO: 128).

Polynucleotides encoding these polypeptides are also provided. Further,

The gene encoding the disclosed cDNA is believed to reside on the X chromosome. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for the X chromosome.

This gene is expressed primarily in human prostate cancer (stage C fraction), subtracted kidney cortex, adult brain and breast, and to a lesser extent in a variety of normal and transformed tissue types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, anti-viral therapies, including cancer and other proliferative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides



are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, kidney, brain and breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types  
5 (e.g., reproductive, neural, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 68 as residues: Ser-28 to Ser-37, Ser-50 to Ser-58. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in various transformed tissues, combined with its similarity to retroviral env proteins indicates that the protein product is useful for the  
15 detection, diagnosis, and treatment of a variety of cancers and other proliferative disorders. This gene may show utility in gene therapy applications in viral prophylaxis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional  
20 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of  
25 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2328 of SEQ ID NO:12, b is an  
30 integer of 15 to 2342, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 3**

The translation product of this gene shares sequence homology with Cytochrome P450 monooxygenase which is thought to be important in NADPH-dependent oxidation of a number of cellular substrates (See Genbank Accession Nos. 5 gi|1185452, gb|AAC50370.1, and gb|AAB87635.1, in addition to Geneseq Accession No. R72378; all information and references available through these accessions are hereby incorporated herein by reference).

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 11 - 27 of the amino acid sequence referenced in 10 Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 28 to 501 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

Included in this invention as a preferred domain is the cytochrome P450 15 cysteine heme-iron ligand signature domain, which was identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Cytochrome P450's [1,2,3] are a group of enzymes involved in the oxidative metabolism of a high number of natural compounds (such as steroids, fatty acids, prostaglandins, leukotrienes, etc) as well as drugs, carcinogens and mutagens. Based on sequence similarities, P450's have been 20 classified into about forty different families [4,5]. P450's are proteins of 400 to 530 amino acids; the only exception is *Bacillus* BM-3 (CYP102) which is a protein of 1048 residues that contains a N-terminal P450 domain followed by a reductase domain. P450's are heme proteins. A conserved cysteine residue in the C- terminal part of P450's is involved in binding the heme iron in the fifth coordination site. From 25 a region around this residue, we developed a ten residue signature specific to P450's. The consensus pattern is as follows: [FW]-[SGNH]-x-[GD]-x-[RHPT]-x-C-[LIVMFAP]-[GAD] [C is the heme iron ligand].

Preferred polypeptides of the invention comprise the following amino acid sequence: FSLGRRHCLG (SEQ ID NO: 129). Polynucleotides encoding these 30 polypeptides are also provided.

Further preferred are polypeptides comprising the cytochrome P450 cysteine heme-iron ligand signature domain of the sequence referenced in Table for this gene,

and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C-terminal to the cytochrome P450 cysteine heme-iron ligand signature domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to the cytochrome P450 cysteine heme-iron ligand signature domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to cytochrome P450 cysteine heme-iron proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with cytochrome P450 cysteine heme-iron proteins. Such activities are known in the art, some of which are described elsewhere herein. The following references were reference above and are hereby incorporated by reference herein: [ 1] Nebert D.W., Gonzalez F.J. *Annu. Rev. Biochem.* 56:945-993(1987). [ 2] Coon M.J., Ding X., Pernecky S.J., Vaz A.D.N. *FASEB J.* 6:669-673(1992). [ 3] Guengerich F.P. *J. Biol. Chem.* 266:10019-10022(1991).[ 4] Nelson D.R., Kamataki T., Waxman D.J., Guengerich F.P., Estrabrook R.W., Feyereisen R., Gonzalez F.J., Coon M.J., Gunsalus I.C., Gotoh O., Okuda K., Nebert D.W. *DNA Cell Biol.* 12:1-51(1993).[ 5] Degtyarenko K.N., Archakov A.I. *FEBS Lett.* 332:1-8(1993).

When tested against Jurket T-cells and U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) promoter element. Thus, it is likely that this gene activates T-cells and myeloid cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in fetal liver spleen, and to a lesser extent in lung, LNCAP prostate cell line, control synovial fibroblasts, human testes tumor, and Hodgkin's lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as  
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatic or developmental diseases and/or disorders, particularly cancer and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential  
10 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, spleen, lung, prostate, testes, and lymphatic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., liver, spleen, developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile,  
15 seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic  
20 epitopes shown in SEQ ID NO: 69 as residues: Leu-30 to Gly-38, Arg-67 to Val-72, Ser-127 to Trp-133, Gly-148 to Phe-154, Thr-171 to Phe-177, Thr-201 to Asp-206, Ser-265 to Pro-273, Glu-283 to Lys-297, Pro-346 to Lys-357, Phe-409 to Glu-418, Glu-423 to Ser-428, Leu-443 to Cys-448. Polynucleotides encoding said polypeptides are also provided.

25 The tissue distribution in fetal liver and homology to cytochrome P450 monooxygenase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In  
30 addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma.

The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the

5 "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g., for treating human immunodeficiency virus infection, cancer, autoimmune diseases

10 and allergy); regulation of hematopoiesis (e.g., for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g., for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g., for treating infections, tumors); hemostatic or thrombolytic activity (e.g., for treating

15 haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g., for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Furthermore, the protein may also be used to determine biological activity, to raise

20 antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1652 of SEQ ID NO:13, b is an

integer of 15 to 1666, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 4

5           The translation product of this gene shares sequence homology with C1qR(P), the human C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro (see, e.g., Genbank Accession number AAB53110.1 (U94333.1); all references available through this accession are hereby incorporated by reference herein. Also see, Immunity 1997 Feb;6(2):119-129).

10           Preferred polypeptides of the invention comprise one or more of the following amino acid sequences:

EHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALSTVRAGAE LRA  
VLALLRAGPGPGXGSKDLLFWVALERRRSHCXLENEPLRGFSWLSSDPGGLE  
SDTLQWVEEPQRSCTARRWV (SEQ ID NO: 130), and/or

15           SRPPVGSSPQLEGDAMPPXRQRYLCKYQFEVLC PAPRPGAASNLSYRAPFQL  
HSAALDFSPPGTEVSALCRGQLPISVT CIADEIGARWDKLSGDVLCPCPGRYL  
RAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPT  
RRPPATATSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMST  
LQMSLQAESKATITPSGVSISKFNSTTSSATPQAFDSSSAVV FIFVSTAVVVLVI  
20           LTMTVLGLVKLCFHESPSSQPRKESMGPPGWRVILKPAALGSSSAHCTNNGV  
KVGDCDLRDRAEGALLAESPLGSSDA (SEQ ID NO: 131). Polynucleotides  
encoding such polypeptides are also provided.

          This gene is believed to reside on chromosome 14. Therefore, polynucleotides  
and polypeptides related to this gene are useful in linkage analysis as markers for  
25           chromosome 14.

          This gene is expressed primarily in chondrosarcoma, smooth muscle tissue,  
bone marrow, chondrosarcoma, fetal tissue (e.g., heart) and to a lesser extent in  
ovarian cancer, adult pulmonary tissues, and brain.

          Therefore, polynucleotides and polypeptides of the invention are useful as  
30           reagents for differential identification of the tissue(s) or cell type(s) present in a  
biological sample and for diagnosis of diseases and conditions which include, but are  
not limited to, chondrosarcoma, immune disorders, ovarian cancer, respiratory and

gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

5 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the

10 disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 70 as residues: Pro-18 to Gly-30. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in bone marrow and homology to C1qR(P) indicates

15 that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation;

20 survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene

25 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues,

30 such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus

erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer (e.g., ovarian, chondrosarcoma), and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.



Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2013 of SEQ ID NO:14, b is an integer of 15 to 2027, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

5

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 5

The translation product of this gene shares sequence homology with human renal dipeptidase, which is a glycosyl-phosphatidylinositol- anchored ectoenzyme thought to be important in the metabolism of dihydro peptide bonds (See Genbank  
 10 Accession No.bbs|148378; dbj|BAA02433.1; gb|AAB59410.1; dbj|BAA02431.1; and Geneseq Accession Nos. W29665 and R30823; all information and references available through these accessions are hereby incorporated herein by reference). Included in this invention as a preferred domain is the renal dipeptidase active site domain, which was identified using the ProSite analysis tool (Swiss Institute of  
 15 Bioinformatics). Renal dipeptidase (rDP) (EC 3.4.13.19), also known as microsomal dipeptidase, is a zinc-dependent metalloenzyme which hydrolyzes a wide range of dipeptides. It is involved in renal metabolism of glutathione and its conjugates. It is a homodimeric disulfide-linked glycoprotein attached to the renal brush border microvilli membrane by a GPI-anchor. A glutamate residue has recently been shown  
 20 [1] to be important for the catalytic activity of rDP. RDP seems to be evolutionary related to hypothetical proteins in the PQQ biosynthesis operons of Acinetobacter calcoaceticus and Klebsiella pneumoniae. The consensus pattern is as follows: [LIVM]-E-G-[GA]-x(2)-[LIVMF]-x(6)-L-x(3)-Y-x(2)-G-[LIVM]-R [E is the active site residue].

25 Preferred polypeptides of the invention comprise the following amino acid sequence: VEGGHSLDNSLSILRTFYMLGVR (SEQ ID NO: 138). Polynucleotides encoding these polypeptides are also provided.

Further preferred are polypeptides comprising the renal dipeptidase active site domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20,  
 30 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the renal dipeptidase active site domain. Alternatively, the additional contiguous

amino acid residues is both N-terminal and C-terminal to the renal dipeptidase active site domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to renal dipeptidase proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with renal dipeptidase proteins. Such activities are known in the art, some of which are described elsewhere herein. The following references were referenced above and are hereby incorporated herein by reference: [ 1 ] Adachi H., Katayama T., Nakazato H., Tsujimoto M. *Biochim. Biophys. Acta* 1163:42-48(1993) and [ 2 ] Rawlings N.D., Barrett A.J. *Meth. Enzymol.* 248:183-228(1995).

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 17 - 33 and 470 - 486 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIb membrane proteins.

When tested against Jurket T-cell cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) and NF-kB (Nuclear Factor kB) pathway. Thus, it is likely that this gene activates T-cells through the Jaks-STAT signal transduction pathway and is involved in the activation of apoptosis. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Similarly, NF-kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB promoter element are used to screen supernatants for such activity.

Preferred polypeptides of the invention comprise the following amino acid sequence: VEGGHS , RYLTLTH SEQ ID NO: 133, CNTPWA SEQ ID NO: 134, APVIFSHS SEQ ID NO: 135, RNVPDD SEQ ID NO: 136, or GLEDVS SEQ ID NO: 137(SEQ ID NO:). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain, and to a lesser extent, in primary dendritic cells, L428 cells, melanocytes, keratinocytes, eosinophils, ovarian tumor, thymus stromal cells, treated bone marrow, and Hodgkins lymphoma, and (to a lesser extent) in a variety of other normal and transformed cell types.

5           Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal, urogenital, or neural disorders, particularly neurodegenerative and/or developmental disorders of the brain, including cancer and other proliferative  
10 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neurological systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types  
15 (e.g., immune, neural, renal, urogenital, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20           Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 71 as residues: Thr-36 to Arg-41, Pro-55 to Pro-60, Pro-67 to Leu-72, Asn-111 to Ser-118, Cys-138 to Asp-144, Asn-290 to Pro-296, Gly-350 to Phe-358, Gly-379 to Glu-384, Gln-399 to Cys-426, Ser-428 to Ser-438. Polynucleotides encoding said polypeptides are also provided.

25           The tissue distribution combined with its homology to the human renal dipeptidase indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney  
30 stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Representative uses are described here and elsewhere herein. Alternatively, the tissue distribution in

brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system.

Moreover, expression within embryonic tissue and other cellular sources marked by proliferating cells combined with the detected GAS and NF-kB biological activity indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2320 of SEQ ID NO:15, b is an integer of 15 to 2334, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

30

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 6**

Preferred polypeptides of the invention comprise the following amino acid

sequence:

TWLRLGSSQIWLGTAPRGPRIHPEQAGLAGAPVKSTSSEESQPGGQCQSSGGA  
 QTLPSLRAAPVAALGSLSSYPDSCPRATTPELCPGAPTLHLADSISSGPVSPPGSS  
 LGPDAWTLCAKHHQAKGMTLGTTPKVLRLQPVSPCWGPKSWRVPGPFQPGR  
 5 RRGESRQQGRGKRRSARSAQSPTGPESAAWPC (SEQ ID NO: 139);  
 TVATACVWAACTGCWARPPVPTWAGCAARCAEDARAGVGDLPATGGAA  
 TGRRALTPAPPRGPCILSPQPWALGLPGAPLPAALPGRARGRPGLPALPALSTL  
 PGCPALDPAGAGTLCPPPGAAEPAGP (SEQ ID NO: 140) or  
 RSGQPGEGSMLRKFSLQRLLSPLDQAQTRWGLALACVAGDKGPPRPWNISSA  
 10 PAHPHVTTPGMETSGGPARDGGLILEREA AFNKPAPGE (SEQ ID NO: 141).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in immune cells (e.g., eosinophils, T-cells, and macrophage), leukemic and lymphoid cells, rectum, colon, and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as  
 15 reagents for differential identification of the tissue(s) or cell type(s) present in a  
 biological sample and for diagnosis of diseases and conditions which include, but are  
 not limited to, inflammatory disorders, particularly immunodeficiency, tumor  
 necrosis, infection, lymphomas, auto-immunities, breast cancer, disorders of the colon  
 and rectum, metastasis, inflammation, anemias (leukemia) and other hematopoietic  
 20 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are  
 useful in providing immunological probes for differential identification of the  
 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,  
 particularly of the immune system, expression of this gene at significantly higher or  
 lower levels is routinely detected in certain tissues or cell types (e.g., immune,  
 25 hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph,  
 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample  
 taken from an individual having such a disorder, relative to the standard gene  
 expression level, i.e., the expression level in healthy tissue or bodily fluid from an  
 individual not having the disorder.

30 The tissue distribution indicates that polynucleotides and polypeptides  
 corresponding to this gene are useful for the diagnosis and treatment of a variety of  
 immune system disorders. Expression of this gene product in a variety of immune

cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2594 of SEQ ID NO:16, b is an integer of 15 to 2608, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of this gene shares sequence homology with the human IgE receptor which is thought to be important in immune regulation, particularly in immune cell aggregation (See Genbank Accession No gi|337418; all references available through this accession are hereby incorporated herein by reference). Moreover, the protein is believed to share structural features to the TM4SF superfamily of proteins.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 73 - 89 and 106 - 122 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with human IgE receptor proteins. Such activities are known in the art, some of which are described elsewhere herein.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

```

RCQRNKDIMSSKPTSHAEVNETIPNPYPSSFMAPGFQQ
PLGSINLENQAQGAQRAQPYGITSPGIFASSQPGQGNIQMINPSVGTAVMNFK
EEAKALGVIQIMVGLMHIGFGIVLCLISFSFREVLFFASTAXIGGYPFWGGLSFI
ISGSLSVSASKELSRCLVKGSLGMNIGRSILAFIGVILLLVDMCINGVXGQDYW
XVLSGKGISATLMIFSXLEFFVACATAHFANQANTTTNMSVLVIPNMYESNPX
TPASSAPPRCNNYSANAPKRKRGISLISWRKTTCKNFLRRCLLLSTMISSL

```



(SEQ ID NO: 142). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in colon.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders and disorders of the digestive tract. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and digestive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 73 as residues: Met-2 to Ser-8, Glu-14 to Ser-23, Leu-39 to Gly-53. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to the human IgE receptor indicates that that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in colon indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of proliferative mechanisms in the digestive tract. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene has homology to a gene of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such

as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1277 of SEQ ID NO:17, b is an integer of 15 to 1291, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 8**

This gene is expressed primarily in T-cells and lymph node.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, particularly immunodeficiencies or inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph,

serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5           The tissue distribution in T-cells and lymph node indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the
- 10   expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).
- 15           Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as
- 20   T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that
- 25   influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue
- 30   markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3115 of SEQ ID NO:18, b is an integer of 15 to 3129, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 9**

This gene is expressed primarily in immune and haemopoietic cells, particularly messangial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or haemopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, haemopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells and hemopoietic cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses

are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3615 of SEQ ID NO:19, b is an integer of 15 to 3629, where both a and b correspond to the positions of nucleotide  
 5 residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 10

The translation product of this gene shares sequence homology with various dehydrogenase and oxidoreductase polypeptides and appears to belong in the alcohol  
 10 dehydrogenase /rybitol dehydrogenase family (see, e.g., Genbank accession number AAD36790.1 (AE001811.1) and CAA68181 (X99908.1); all references available through this accession are hereby incorporated by reference herein.).

Preferred polypeptides of the invention comprise the following amino acid sequence:

15 MGRLDGKVIILTAAAQGIGQAAALAFAREGAKVIATDINESKLQELEYPGIQ  
 TRVLDVTKKKQIDQFANEVERLDVLFNVAGFVHHGTVLDCEEKDWDFSMNL  
 NVRNVMYLMIKAFLPKMLAQKSGNIINMSSVASSVKGVVNRCVYSTTKAAV  
 IGLTKSVAADFIQQGIRCNCVCPGTVDTPSLQERIQARGNPPEARNDFLKRQK  
 TGRFATAEEIAMLCVYLASDESAYVTGNPVIIDGGWSL (SEQ ID NO: 143).

20 Also provided are fragments thereof having dehydrogenase activity and polypeptides comprising at least 30 residues of the foregoing amino acid sequence. Polynucleotides encoding such polypeptides are also provided.

This gene is expressed primarily in fetal tissue (e.g., liver, spleen, lung), gall  
 25 bladder, heart, bone marrow and to a lesser extent in smooth muscle, and parathyroid tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, cardiovascular and developmental disorders. Similarly,  
 30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and

fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a  
5 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 76 as residues: Pro-78 to Gln-85, Arg-87 to Arg-94, Asp-96 to Gly-104. Polynucleotides encoding said polypeptides are also provided.

10 The tissue distribution in bone marrow indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene  
15 product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

20 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as  
25 T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that  
30 influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,

and in the differentiation and/or proliferation of various cell types. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically



excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1130 of SEQ ID NO:20, b is an integer of 15 to 1144, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 11

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 115 - 131 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 132 to 152 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GTIGLYWVGSII MSVVVFVPGNIVGKYGTRICPAFFLSIPYTCLPVWAGFRIYN  
 QPSENYNYPSKVIQEAQAKDLLRRPFDLMLVVCLLLATGFCLFRGLIALDCPS  
 ELCRLYTQFQEPYLKDPAAYPKIQMLAYMFYSVPYFVTALYGLVVPGCSWM  
 PDITLIHAGGLAQAQFSHIGASLHARTAYVYRVPEEAKILFLALNIAYGVL PQL  
 LAYRCIYKPEFFIKTKAEKVE (SEQ ID NO: 144). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in osteoclastoma, and to a lesser extent, in other human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal diseases and/or disorders, particularly osteoclastoma and osteoporosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the

tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., skeletal, immune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 77 as residues: Thr-32 to Lys-40, Lys-146 to Glu-152. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in osteoclastoma indicates a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g., arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Representative uses are described here and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1429 of SEQ ID NO:21, b is an integer of 15 to 1443, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 12

The polypeptide of this gene has been determined to have transmembrane domains at about amino acid position 1-23 and 149-167 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIb membrane  
10 proteins.

Preferred polypeptides of the invention comprise the following amino acid sequence:

MSNHDPRGCTRRRAQKPLAIQPRLFHASAPDEGTQGTLKGTQKGGCILVQCQ  
SEGGAAGAWTGPPSPARDRRVRPPGTKAQRLERRRHVPRLHGLGVGGCEVR  
15 TGIVARISGSTPWAGGKPLGLHGAMGEAGAGDTGCCAKGSPAAPLPAEGRG  
QGAGPGGLVGRGERRDQQTLLGMAEDTGXSPSRPSAPAPRAPVPARQPLPRA  
RLGAATAISKSRSSRVAPALAAASASSHQR (SEQ ID NO: 145). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in neutrophils, haemopoietic cells, tymus  
20 tumor, osteosarcoma, synovial sarcoma, B-cell lymphoma, dendritic cells, pineal gland, brain, prostate and to a lesser extent in other tissues, including cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
25 not limited to, immune and haemopoietic disorders, particularly neutropenia or neutrophilia, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic systems, expression of  
30 this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, haemopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 78 as residues: Ser-23 to Ala-32, Gly-40 to Glu-47. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells (e.g., neutrophils) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in tonsils also indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell

lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses). The tissue distribution in pineal gland and brain indicates polynucleotides and

5 polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection,

10 treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder,

15 learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation,

20 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show

25 utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in thymus tumor, B-cell lymphoma, osteosarcoma, and synovial sarcoma indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of these and related diseases.

Many polynucleotide sequences, such as EST sequences, are publicly

30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1039 of SEQ ID NO:22, b is an integer of 15 to 1053, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The translation product of this gene shares sequence homology with lymphoblastic leukaemia antigen, which is thought to be important in cancers including leukemia.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 46 - 62 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 63 to 69 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

This gene is expressed primarily in lung, and infant adrenal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the pulmonary and endocrine system, including cancers and developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary and endocrine systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pulmonary, endocrine, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in adrenal gland indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", and "Binding Activity" sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes.

Similarly, expression within fetal tissues and other cellular sources marked by proliferating cells, combined with the homology to the human lymphoblastic leukaemia antigen, indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 727 of SEQ ID NO:23, b is an integer of 15 to 741, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

## 5    **FEATURES OF PROTEIN ENCODED BY GENE NO: 14**

This gene is expressed primarily in human tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
10    not limited to, immune or hematopoietic diseases and/or disorders, particularly leukemia and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at  
15    significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from  
20    an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 80 as residues: Gly-33 to Arg-40, Ser-106 to Met-112, Ala-154 to Gly-163. Polynucleotides encoding said polypeptides are also provided.

25    The tissue distribution in tonsils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product  
30    indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen



presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory conditions such as inflammatory bowel disease, sepsis, acne, and psoriasis and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 932 of SEQ ID NO:24, b is an integer of 15 to 946, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 15**

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 49 - 65 and 141 - 157 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

5 STXTXTIGXAGTPAGTGPEFPGRPTRPGEXPVDFSKQYSASWMCLSLAALA  
CSAGDTWASEVGPVLSKSSPRLITTWEKVPVGTNGGVTVVGLVSSLLGGTFV  
GIAYFLTQLIFVNDLDISAPQWPIAFGGLAGLLGSIVDSYLGATMQYTGLDES  
TGMVVNSPTNXARHIAGKPILDNNAVNLFSSVLIALLLPTAAWGFWPRG  
(SEQ ID NO: 146). Polynucleotides encoding these polypeptides are also provided.

10 This gene is expressed primarily in colon, brain, and to a lesser extent, in epiglottis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
15 not limited to, diseases and/or disorders of the central nervous system and gastrointestinal or digestive tract. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and central nervous system, expression of  
20 this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointesinal, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or  
25 bodily fluid from an individual not having the disorder.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11,  
30 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia,

obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. The protein product of this gene may also be useful for the detection, treatment, or prevention of a variety of gastrointestinal and digestive tract disorders, particularly proliferative disorders, such as ulcers and cancers. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 817 of SEQ ID NO:25, b is an integer of 15 to 831, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

## **25 FEATURES OF PROTEIN ENCODED BY GENE NO: 16**

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 21-39 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 40-41 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

Preferred polypeptides comprise the following amino acid sequence:

MSQRAG

RRPGGWNPSLSVVEVCRGCRGTGPLPWGASLFPCASPLFPLPLNRRGDVHG  
TLGGRMLNRVECRDGVAAAWLCLHDAAAIRGAVGRCPMWTQPTHWVLLL  
CWALHFYCR (SEQ ID NO: 147) Also preferred are the polynucleotides encoding

5 these polypeptides.

This gene is expressed primarily in tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
10 not limited to, immune disorders, particularly tonsillitis and tonsillar neoplasms. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is  
15 routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, saliva, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the  
20 disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of  
25 potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene  
30 product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel

disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1280 of SEQ ID NO:26, b is an integer of 15 to 1294, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

## **25 FEATURES OF PROTEIN ENCODED BY GENE NO: 17**

This gene is expressed primarily in fetal tissue (e.g., bone, liver, spleen), smooth muscle, chondrosarcoma, osteoblasts, osteosarcoma, and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of developing and growing organs and tissues, bone disease, osteosarcoma, and other cancers. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., fetal tissue, bone, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, bone marrow, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10       The tissue distribution in osteosarcoma, osteoblasts, and chondrosarcoma indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of bone disease and diseases of the skeletal system. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

20       Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in

modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue  
5 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly  
10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or  
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1642 of SEQ ID NO:27, b is an integer of 15 to 1656, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

## 20 FEATURES OF PROTEIN ENCODED BY GENE NO: 18

This gene is expressed primarily in liver, fetal liver, and to a lesser extent in bone marrow stromal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a  
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatic dysfunction, immune disorders, and disease of the hemopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,  
30 particularly of the liver and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, liver, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,

urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 84 as residues: Glu-44 to Asp-50. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in liver and fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment  
10 of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. The  
15 tissue distribution in bone marrow indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product  
20 indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 25 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as  
30 T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus



erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful

5 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

10 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of

15 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1336 of SEQ ID NO:28, b is an

20 integer of 15 to 1350, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 19**

The translation product of this gene shares sequence homology with vacuolar

25 proton-ATPase subunit M9.2 (see, e.g., Genbank accession numbers CAA75571 (Y15286.1); all references available through this accession are hereby incorporated by reference herein.). Preferred polypeptide encoded by this gene comprise the following amino acid sequence:

MTAHSFALPVIIFTTFWGLVGIAGPWFVPKGPNRGVITMLVATAVCCYLFWL

30 IAILAQLNPLFGPQLKNETIWYVRFLWE (SEQ ID NO: 148) and

AQRAARLGTRAPAAPAARPCILPGHPAPGHDGALIRPPGHHLHHVLGPRRHR

GPWFVPKGPNRGVITMLVATAVCCYLFWLIAILAQLNPLFGPQLKNETIWYV

RFLWE (SEQ ID NO: 149). Polynucleotides encoding such polypeptides are also provided.

This gene is expressed primarily in infant brain, pancreas islet cell tumor, ovary tumors, immune cells (e.g., T-cells), normal cerebellum, endometrial tumor  
5 tissues and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodevelopmental disorders, endocrine system disorders, disorders  
10 of the immune system, and ovarian cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, ovaries, immune system, expression of this gene at significantly higher or lower levels is routinely  
15 detected in certain tissues or cell types (e.g., immune, neural, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution and homology to vacuolar proton-ATPase subunit M9.2 indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodevelopmental disorders. The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative  
25 disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis,  
30 encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive

disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5           Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The tissue distribution in immune cells (e.g., T-cells) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses  
10 are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine  
15 production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

          Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such  
20 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus  
25 erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue  
30 distribution in endocrine tissues such as the pancreas indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or

prevention of various endocrine disorders and cancers. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", and "Binding Activity" sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of

5 Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to

10 isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

15 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1752 of SEQ ID NO:29, b is an integer of 15 to 1766, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

## 25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 20**

This gene is expressed primarily in placenta, induced endothelial cells, immune cells (e.g, T-cells, B-cells, leukocytes), brain, fetal tissue, epididymus, lung, lung cancer, thyroid tumor and to a lesser extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as

30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal developmental disorders, immune disorders, cancer of the lungs,

thyroid, and cancer, in general. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

5 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the

10 disorder.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative

15 Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction,

20 aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

25 Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The tissue distribution in immune cells (e.g., T-cells, B-cells, and leukocytes) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.

30 Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the

proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

5           Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as  
10 T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that  
15 influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates  
20 this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell  
25 differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of  
30 potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of  
5 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation.

The tissue distribution in endothelial cells indicates that polynucleotides and  
10 polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the vasculature. Elevated expression of this gene product by endothelial cells indicates that it may play vital roles in the regulation of endothelial cell function; secretion; proliferation; or angiogenesis. Alternately, this may represent a gene product expressed by the endothelium and transported to distant sites of action  
15 on a variety of target organs. Expression of this gene product by hematopoietic cells also indicates involvement in the proliferation; survival; activation; or differentiation of all blood cell lineages. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a  
20 nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
25 related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general  
30 formula of a-b, where a is any integer between 1 to 2776 of SEQ ID NO:30, b is an integer of 15 to 2790, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 21**

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 1-29 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

Preferred polypeptides comprise the following amino acid sequence:

MTLEEHRDRPRLGMCMCVCACVYACMLMHVCVHACLCVCVCVCVEPWSS  
R

QSKDTGGWHMEEQVTPPSLAQLKSGQVRGEMGEGRGEKGEEALTGGAEAL  
SLLGRRSPSTPLFLDREDKQAKDARNLSSTVAPDF (SEQ ID NO: 150). Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in tonsils and activated monocytes and to a lesser extent in activated neutrophils and anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 87 as residues: Thr-27 to Arg-33, Gly-37 to Ser-42, Pro-52 to Arg-72. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells (e.g., T-cells, neutrophils) indicates polynucleotides and polypeptides corresponding to this gene are useful for the



diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or  
5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1403 of SEQ ID NO:31, b is an integer of 15 to 1417, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

#### 10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 22**

This gene is expressed primarily in CD34 positive cells (Cord Blood) and resting T-cells and to a lesser extent in anergic T-cells and neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a  
15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at  
20 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the  
25 disorder.

The tissue distribution in immune cells (e.g., T-cells, neutrophils) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in  
30 Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene  
5 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues,  
10 such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits  
15 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their  
20 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
25 related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general  
30 formula of a-b, where a is any integer between 1 to 1892 of SEQ ID NO:32, b is an integer of 15 to 1906, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 23**

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 33 - 49 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 50 to 62 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

HEKILTPIWPSSTDLEKPHLEMLFLNVILFSLTVFTLISTAHTLDRAVRSDWLLL  
VLIYACLEELIPELIFNLYCQGNATLFF (SEQ ID NO: 151). Polynucleotides

encoding these polypeptides are also provided.

This gene is expressed primarily in activated T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic diseases and/or disorders, particularly inflammatory conditions or immunodeficiencies such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of

immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or  
5 activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene  
10 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues,  
15 such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits  
20 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their  
25 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
30 related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 529 of SEQ ID NO:33, b is an integer of 15 to 543, where both a and b correspond to the positions of nucleotide  
5 residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 24**

This gene is expressed primarily in activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as  
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic diseases and/or disorders, particularly inflammatory or immunodeficiency disorders, such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological  
15 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal  
20 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a  
25 variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells.  
30 This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1438 of SEQ ID NO:34, b is an integer of 15 to 1452, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 25**

This gene is expressed primarily in multiple sclerosis tissue, immune cells (e.g., T-cells and dendritic cells), brain, uterus, ovary, stomach, placenta, and fetal tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, multiple sclerosis, disorders of the fetus and female reproductive system, immune disorders, particularly immunodeficiency, tumor necrosis, infection,

lymphomas, auto-immunities, cancer, metastasis, wound healing, inflammation, anemias (leukemia) and other hematopoietic disorders, in addition to developmental or proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential  
5 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, developmental, reproductive system, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid,  
10 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic  
15 epitopes shown in SEQ ID NO: 91 as residues: Met-1 to Lys-6. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells (e.g., T-cells, dendritic cells) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses  
20 are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine  
25 production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such  
30 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues,



such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that

5 influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue distribution in female reproductive organs indicates that polynucleotides and

10 polypeptides corresponding to this gene are useful for treating female infertility and cancers including but not limited to ovaries and uterus. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein

15 products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. The tissue distribution in brain indicates polynucleotides and

20 polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection,

25 treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder,

30 learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2894 of SEQ ID NO:35, b is an integer of 15 to 2908, where both a and b correspond to the positions of nucleotide  
10 residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 26**

The gene encoding the disclosed cDNA is believed to reside on chromosome 20. Accordingly, polynucleotides related to this invention are useful as a marker in  
15 linkage analysis for chromosome 20.

This gene is expressed primarily in fetal and developing tissues, tumors of male and female reproductive tissue (e.g., ovary and testes), and immune cells (e.g., T-cells).

Therefore, polynucleotides and polypeptides of the invention are useful as  
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and growth disorders, disorders of the immune system, disorders and cancers of ovaries and testes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for  
25 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal, reproductive, or developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, differentiating, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid,  
30 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 92 as residues: Val-57 to Ala-63. Polynucleotides  
5 encoding said polypeptides are also provided.

The tissue distribution in immune cells (e.g., T-cells) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in  
10 Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the  
15 treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel  
20 disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's  
25 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue  
30 distribution in ovaries and testes indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of diseases of the reproductive organs, including but not limited to ovarian and testicular cancer. The

expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative  
5 uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent  
10 of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the  
15 polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in  
20 modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their  
25 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
30 related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 939 of SEQ ID NO:36, b is an integer of 15 to 953, where both a and b correspond to the positions of nucleotide  
5 residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 27**

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as  
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and haemopoietic disorders, particularly immunodeficiencies, such as AIDS, or inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for  
15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, haemopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal  
20 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of  
25 immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene  
30 product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3850 of SEQ ID NO:37, b is an integer of 15 to 3864, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 28**

In another embodiment, polypeptides comprising the amino acid sequence of

the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PANKAGAAIEAGIGISLMVLSPWACLFVVFPPYIQSSLRSDKHLQLSNILPTPS

5 HHIHLPASICIQLRAGN (SEQ ID NO: 152). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain (early stage human brain and cerebellum) and immune cells (activated neutrophils, activated T-cells, neutrophils and dendritic cells) and, to a lesser extent in other tissues.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural or immune diseases and/or disorders, particularly inflammatory disorders, in addition to cancer and other proliferative disorders. Similarly,  
15 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain or the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural,  
20 and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the  
30 detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses ,



autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system.

- 5 Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in a variety of hematopoietic derived cells (T-cells, neutrophils, etc.) indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of
- 10 potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene
- 15 product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages,
- 20 and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
- 25 immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
- 30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1397 of SEQ ID NO:38, b is an integer of 15 to 1411, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 29

When tested against Jurket T-cell cell lines, supernatants removed from cells containing this gene activated the NF-kB (Nuclear Factor kB) promoter element. Thus, it is likely that this gene activates T-cells. NF-kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB promoter element are used to  
10 screen supernatants for such activity.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 5 - 21 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 22 to  
15 40 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by  
20 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

AGSPAGTGPEFPGRPTRPISTHVFEYECICKIPRFMCEYVLLLYIVLLCNRSYA  
VFTQCVLRSSPIDSSRNAVLL (SEQ ID NO: 153). Polynucleotides encoding these polypeptides are also provided.

25 This gene is expressed primarily in colon, synovium, chondrosarcoma and to a lesser extent in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
30 not limited to, immune, gastrointestinal, or skeletal diseases and/or disorders, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and digestive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., lymph, skeletal, gastrointestinal, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g., arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid).

The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities, particularly NF $\kappa$ -B activation. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g., for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g., for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g., for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g., for treating infections, tumors); hemostatic or thrombolytic activity (e.g., for treating

haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g., for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures.

- 5 Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is  
15 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1168 of SEQ ID NO:39, b is an integer of 15 to 1182, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

20

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 30**

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

- 25 This gene is expressed primarily in immune cells (e.g., activated T-cells), testes, fetal, spleen, and to a lesser extent in colon tumor, teratocarcinoma cells, brain and number of other tissues and cell types.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a  
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders related to T-cell dysfunction, such as immunodeficiencies or inflammatory conditions, in addition to neural, or

reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

5 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression

10 level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 96 as residues: Pro-25 to Arg-38. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in activated T-cells and spleen indicates that

15 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune related disorders and diseases, including hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases; immunodeficiency diseases, such as acquired immunodeficiency syndrome,

20 autoimmunity, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma; infections, and other inflammatory diseases and complications. Additionally, the tissue distribution in T-cells and spleen indicates polynucleotides and polypeptides corresponding to this gene are useful for the

25 diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

30 blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of

degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues.

The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. In addition, the secreted protein can be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities although no evidence for any is provided in the specification. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g., for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of haematopoiesis (e.g., for treating anaemia or as adjunct to chemotherapy); stimulation of growth of bone, cartilage, tendons, ligaments and/or nerves (e.g., for treating wounds, periodontal disease, neurological diseases stroke, fibrosis); inhibition or stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g., for treating infections, tumours); haemostatic or thrombolytic activity (e.g., for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g., for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative disease; for regulation of metabolism, behaviour, and many others. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2443 of SEQ ID NO:40, b is an integer of 15 to 2457, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with epidermal growth factor which is thought to be important in the growth and proliferation of epidermal cells, fibroblasts and number of other cell types and tissues.

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, particularly autoimmunities or connective tissue diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and connective tissue systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, epithelial, endothelial, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14,



16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or  
5 other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such  
10 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus  
15 erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,  
20 and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,  
25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 1833 of SEQ ID NO:41, b is an integer of 15 to 1847, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 32

This gene is expressed primarily in kidneys, tonsils and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
10 not limited to, renal disorders and neoplasms, tonsilitis and immune disorders, particularly infections. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal and immune systems, expression of this gene  
15 at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, urogenital, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in  
20 healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of renal conditions such as acute renal failure, kidney fibrosis and kidney tubule regeneration and neoplasms. Conditioned media, generated from the transient  
25 expression of this gene in CHO cells has been shown to activate a IFN $\gamma$ -reponsive element (GAS) in a B cell line (Reh). The same conditioned media had no effect on T-cell (Jurkat) and pro-monocyte (U937) derived cell lines, suggesting that the protein product of this gene may exhibit IFN $\gamma$ -like activity in a (B)cell-specific manner. This experimental data in conjunction with expression on dendritic cells  
30 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g., AIDS), immuno-suppressive

conditions (transplantation) and hematopoietic disorders. In addition this gene product is applicable in conditions of general microbial infection, inflammation or cancer.

Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents  
 5 that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
 10 related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or  
 15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2583 of SEQ ID NO:42, b is an integer of 15 to 2597, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 33**

20 The translation product of this gene was shown to have homology to the classII MHC transactivator CIITA of Mus musculus (See, e.g., Genbank Accession No gi|1870520 and AAB48859.1; all references available through this accession are hereby incorporated by reference herein.), which is thought to regulate MHC class II  
 25 gene expression in B lymphocytes via direct interaction with the MHC class II-specific transcription factors. Furthermore, the CIITA protein is thought to play an indirect role in reducing tumorigenicity and inducing long-term tumor immunity.

Preferred polypeptides of the invention comprise the following amino acid sequence:

MPSGMSAAVPISGLLDLSHNSISQESALYLLETLPSCPRVREASVNLGSEQSF  
 30 RIHFSREDQAGKTLRLSECSFRPEHVSRLATGLSKSLQLTELTLTQCCLGQKQL  
 AILLSLVGRPAGLFSRLRVQEPWADRARVLSLLEVCAQASGSVTEISISSETQQQL  
 CVQLEFPRQEENPEAVALRLAHCDLGAHHSLLXGQLMETCARLXQLSLSQV

NLCEDDDASSLLLSQSLLSLSELKTFRLTSSCVSTEGLAHLASGLGHCHHLEEL  
 DLSNNQFDEEGTKALMRALEGKWMLKRLDLSHLLLSSTLALLTHRLSQMT  
 CLQSLRLNRNSIGDVGCCHLSEALRAATSLEELDLSHNQIGDAGVQHLATILP  
 GLPELRKIDLSGNSISSAGGVQLAESLVLCRRLEELMLGCNALGDPTALGLAQ  
 5 ELPQHLRVLHLPFSLGPGGALSARPWMDPPIWKRSAWRKTTWLEGSCVS  
 VWSSRCSDR (SEQ ID NO: 154); or  
 HQLSRGSAVGRVSRSLQAPGGVDAWLQCPGGSHSPGAGSGAAPAPEGPTPTI  
 QPSGPRWGPEPGQALDGSPHLEEISLAENNLAGGVLRFCMELPLLQIDLVSC  
 KIDNQTAKLLTSSFTSCPALEVILLSWNLLGDEAAELAQVLPQMGRLEKRV  
 10 LE  
 KNQITALGAWLLAEGLAQGSSIQVIRLWNNPIPCDMAQHLKSQEPRLDFAFFD  
 NQPQAPWGT (SEQ ID NO: 155). Polynucleotides encoding these polypeptides are  
 also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome  
 15 16. Accordingly, polynucleotides related to this invention are useful as a marker in  
 linkage analysis for chromosome 16.

The polypeptide of this gene has been determined to have transmembrane  
 domains at about amino acid positions 1-47 and 24-47 of the amino acid sequence  
 referenced in Table 1 for this gene. Based upon these characteristics, it is believed  
 20 that the protein product of this gene shares structural features to type IIIa membrane  
 proteins.

This gene is expressed primarily in immune cells (e.g., eosinophils, T-cells,  
 dendritic) and other cell types of hematopoietic origin and to a lesser extent in ovary  
 tumor and heart.

25 Therefore, polynucleotides and polypeptides of the invention are useful as  
 reagents for differential identification of the tissue(s) or cell type(s) present in a  
 biological sample and for diagnosis of diseases and conditions which include, but are  
 not limited to, immune disorders, particularly immunodeficiency, tumor necrosis,  
 infection, lymphomas, auto-immunities, cancer, metastasis, inflammation, anemias  
 30 (leukemia) and other hematopoietic disorders. Similarly, polypeptides and antibodies  
 directed to these polypeptides are useful in providing immunological probes for  
 differential identification of the tissue(s) or cell type(s). For a number of disorders of

the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another  
5 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells (T-cells, dendritic, and eosinophils) indicates polynucleotides and polypeptides corresponding to this gene are useful for  
10 the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including  
15 blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for  
20 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders,  
25 such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful  
30 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5           Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
- 10           cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3102 of SEQ ID NO:43, b is an integer of 15 to 3116, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 34**

Polynucleotides of the invention do not comprise the nucleic acid sequence shown as Genbank accession no.gb|Z85986|HS108K11, which is hereby incorporated herein by reference.

- 20           The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

This gene is expressed primarily in smooth muscle and cells of hematopoietic origin.

- 25           Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular disorders, particularly heart disease, vasculitis, atherosclerosis, in addition to immune disorders, such as immunodeficiency, auto-
- 30           immunities, cancer, metastasis, anemias (leukemia) and other hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the

tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the [insert system where a related disease state is likely, e.g., immune], expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, endothelial, muscle, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10           The tissue expression in hematopoietic tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

20           Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,

and in the differentiation and/or proliferation of various cell types. In addition, the expression in smooth muscle might indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cardiovascular and disorders such as atherosclerosis, restenosis, stoke, angina, thrombosis

5 hypertension, inflammation and vascular wound healing. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

10 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3446 of SEQ ID NO:44, b is an integer of 15 to 3460, where both a and b correspond to the positions of nucleotide

20 residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 35**

The translation product of this gene was shown to have homology to the human cytochrome c-like polypeptide from lung adenocarcinoma A549 (See

25 Genbank Accession No.bbs|175350) which is thought to be involved in metabolic processes, specific to tumors or transformed cells.

Preferred polypeptides of the invention comprise the following amino acid sequence:

EKLFCFEMLLICKFSPNSVPPETCAILNQGLMDLGLCRMCLGNNMFAGSMLG

30 KSHRHSPFSINQRHNAIRKAAGTPAQKSLGIVQVS PN SEQ ID NO: 156,  
GCAGCALVTICLQAVCLVKAIALHSRLTRDTMHCGRPQGPLPRKAWVLSRF  
PPTETA (SEQ ID NO: 157);



PETQCTAEGRRDPCPEKPGYCPGFPQLRQPEIWPRGKGKTLHPPARHM (SEQ ID NO: 158); SEIGENRP (SEQ ID NO: 159); HDTDSFAH (SEQ ID NO: 160); or ALRKAAG (SEQ ID NO: 161). Polynucleotides encoding these polypeptides are also provided. Polynucleotides of the invention do not comprise the nucleic acid sequence  
5 shown as Genbank accession no.gb|AC004706|AC004706, which is hereby incorporated herein by reference.

This gene is expressed primarily in fetal tissue, lung, melanocyte, retina, brain, T-cell lymphoma, and to a lesser extent, in other tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as  
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hepatic, pulmonary, developmental, or growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell  
15 type(s). For a number of disorders of the above tissues or cells, particularly of the immune and pulmonary systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, pulmonary, hematopoietic, developing, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial  
20 fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of  
25 neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,  
30 trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive

compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

5           Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The tissue distribution in T-cell lymphoma indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the  
10   "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or  
15   other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

          Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such  
20   as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus  
25   erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,  
30   and in the differentiation and/or proliferation of various cell types. Moreover, the expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may

show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions  
5 involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of  
10 potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types  
15 of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new  
20 insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,  
25 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2608 of SEQ ID NO:45, b is an integer of 15 to 2622, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 36

The translation product of this gene was shown to have homology to the human protein kinase C substrate 80K-H (See Genbank Accession No. P14314), which is important in the regulation of various signal transduction pathways. Polynucleotides of the invention do not comprise the nucleic acid sequence shown as  
10 Genbank accession no. gb|G21007|G21007, which is hereby incorporated herein by reference.

This gene is expressed primarily in breast and prostate cancer, retina, ovary, parathyroid tumor, fetal tissue and to a lesser extent in ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as  
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, endocrine, or ocular disorders, particular breast or other cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the  
20 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and endocrine systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, endocrine, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, breast milk, aqueous humor, vitreous humor, serum,  
25 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic  
30 epitopes shown in SEQ ID NO: 102 as residues: Gln-14 to Val-20, Arg-28 to Trp-35. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in breast cancer cell lines, combined with the homology to a PKC substrate indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis, treatment, and/or prevention of a variety of tumors, particularly of the breast or other neoplasms. Alternatively, the tissue  
5 distribution in ovary and parathyroid tumors indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid  
10 (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular  
15 division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern  
20 formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of  
25 potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types  
30 of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in

modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in immune cells (e.g., germinal B-cells) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the

protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in parathyroid tumor, prostate cancer and breast cancer tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment these and related disorders.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1970 of SEQ ID NO:46, b is an integer of 15 to 1984, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

## 20 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

Preferred polypeptides of the invention comprise the following amino acid sequence:

MRGPVCGFSLVEMLLALALGLMLILGVTQIALSSRTTYASQSAASLLQDDAR  
FALGKLIQEIRQAGMFGCLSAASISNAPAGFDRPIGWSTTGSSRSLTLVTADVG  
25 EGGSKPDWTVLSDCTGSAHAYVGSPPAANARANPLPTCAKLT (SEQ ID NO:  
162); or  
MGYYLSRSRQAGMVLLISLVFLLLLALLGVSSMQGAISQEKITGSLRQRNQSF  
QQAESGLRLGESLVQASGFALRPCHSTAACAPPAESVSVVGPGTNPVSTVTWI  
GMKDG VYGIQNLGPGTGLVNSRQRPRPRISA (SEQ ID NO: 163).

30 Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in cord blood, ovary, tumors of the parathyroid, testes, and pancreas, and to a lesser extent in fetal tissue, retina, brain, colon, endometrial stromal, HL-60 cells, and many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as  
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, or developmental disorders, cancer of the ovaries and endocrine system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential  
10 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, reproductive, developmental, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid,  
15 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic  
20 epitopes shown in SEQ ID NO: 103 as residues: Lys-29 to Ser-38, Ser-55 to Trp-61, Gln-63 to Ser-69. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in endocrine tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's Disease,  
25 Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothalamus, and testes. The tissue distribution in immune cells (e.g., neutrophils and T-cells) indicates polynucleotides and polypeptides corresponding to this gene are  
30 useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein.



Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may

also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation.

10       The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

20       Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30       Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or  
5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1973 of SEQ ID NO:47, b is an integer of 15 to 1987, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

#### 10 FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 1-24 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

15 This gene is expressed primarily in activated monocytes, dendritic cells, and cancerous ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are  
20 not limited to, ovarian cancer, immune or hematopoietic disorders, particularly immunodeficiencies or inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene  
25 at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily  
30 fluid from an individual not having the disorder.

The tissue distribution in monocytes and dendritic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow  
5 reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. The tissue distribution in dendritic cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a  
10 variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene  
15 product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for  
20 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders,  
25 such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful  
30 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue distribution in ovarian cancer tissue indicates that polynucleotides and polypeptides

corresponding to this gene are useful for the treatment and diagnosis of ovarian, as well as, other cancers:

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2099 of SEQ ID NO:48, b is an integer of 15 to 2113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 39**

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 36-57 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type II membrane proteins.

Preferred polypeptides of the invention comprise the following amino acid sequence:

ENESTKEPSLLQYLCVQSPAGLNGFNVLLSGSQTPTTVGPSSGQLPSFSVPCM  
VLPSPPLGPFVLYSPAMPGVSSSTLGALPNTGPVNFSLPGLGSIAQLLVGPTA  
VVNPKSSTLPSADPQLQSQPSLNLSPVMSRSHSVVQQPESPVYVGHPVSVVKL  
HQSPVPVTPKSIQRTTHRETFKTPGSLGDPVLKRRERNNHETPARPRGD (SEQ

ID NO: 164). Polynucleotides encoding such polypeptides are also provided.

This gene is expressed primarily in 8 week-old embryo, fetal lung, testes, colon, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or

activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5           Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as
- 10   T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that
- 15   influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. endothelium, and T-cell lymphoma.
- 20           Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and immune defects, cancer, T-cell lymphoma, and developmental disorders. Similarly, polypeptides and antibodies directed to these
- 25   polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,
- 30   urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cell lymphoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of developmental and pulmonary defects and neoplasms of blood, reproductive and other organs. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The tissue distribution in testes tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of male reproductive and endocrine disorders. It may also prove to be valuable in the

diagnosis and treatment of testicular cancer, as well as cancers of other tissues where expression has been observed. The expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,  
5 and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

10       Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have  
15 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue  
20 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the  
25 protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30       Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of



the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general  
5 formula of a-b, where a is any integer between 1 to 3451 of SEQ ID NO:49, b is an integer of 15 to 3465, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 40**

10 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 145-171 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 168-282 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia  
15 membrane proteins.

This gene is expressed primarily in T helper cells, breast cancer, kidney, fetal tissue and to a lesser extent in thymus and cells from some other hemopoietic and endocrine sources.

Therefore, polynucleotides and polypeptides of the invention are useful as  
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory conditions, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of  
25 disorders of the above tissues or cells, particularly of the hemopoietic and lymphoid systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the  
30 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 106 as residues: Pro-135 to Ile-145, Trp-173 to Gly-188, Pro-199 to Gln-219, Ser-225 to Ala-237, Pro-240 to Gly-253, Ser-262 to Gly-275. Polynucleotides encoding said polypeptides are also provided.

- 5           The tissue distribution in breast cancer tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neoplasms of breast and other organs. The tissue distribution in immune cells (e.g., T-helper cells) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.
- 10   Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation
- 15   of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such
- 20   as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus
- 25   erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,
- 30   and in the differentiation and/or proliferation of various cell types. The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal

tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Moreover, the expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1223 of SEQ ID NO:50, b is an integer of 15 to 1237, where both a and b correspond to the positions of nucleotide  
10 residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 41**

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in  
15 linkage analysis for chromosome 16.

This gene is expressed primarily in adipose tissue, brain, Hodgkin's lymphoma, and to a lesser extent in fetal tissue colon tumor, synovium, salivary gland, immune cells (e.g., neutrophils), and other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as  
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders, immune, or metabolic disorders, particularly diseases or disorders of adipose tissue or Hodgkin's lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing  
25 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adipose tissues and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, adipose, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,  
30 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 107 as residues: Arg-54 to Leu-60, Ala-73 to Gly-78.

5 Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of obesity, or other metabolic disorders, such as Tay-Sach's Disease, phenylketonuria, galactosemia, hyperlipidemias, porphyrias, and Hurler's syndrome, in addition to  
10 various immune disorders and neoplasia. The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example  
15 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms,  
20 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

25 Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The tissue distribution in immune cells (e.g., neutrophils) indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses  
30 are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation;

survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5           Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as
- 10   T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that
- 15   influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. In addition, the expression of this gene product in synovium would suggest a role in the detection and
- 20   treatment of disorders and conditions afflicting the skeletal system, in particular osteoporosis, bone cancer, connective tissue disorders (e.g., arthritis, trauma, tendonitis, chrondomalacia and inflammation). The protein is also useful in the diagnosis or treatment of various autoimmune disorders (i.e., rheumatoid arthritis, lupus, scleroderma, and dermatomyositis), dwarfism, spinal deformation, joint
- 25   abnormalities, and chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid, etc.). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional
- 30   supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1383 of SEQ ID NO:51, b is an integer of 15 to 1397, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 42

Polynucleotides of the invention do not comprise the nucleic acid sequence shown as Genbank accession no.gb|G15452|G15452, which is hereby incorporated herein by reference.

Preferred polypeptides of the invention comprise the following amino acid sequence:

RHERHEYRRALDHEEEALSSGSVQEAEAMLDEPQEQAEGSLTVYVISEHSSLL  
 PQDMMSYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQAMSLTEDVLAAALAD  
 20 HLPEDKWSAEKRRPLKSSLGYEITFSLNPDPKSHDVYWDIEGAVRRYVQPFL  
 NALGAAGNFSVDSQILYYAMLGVNPRFDSASSSYLDMHSLPHVINPVESRL  
 GSSAASLYPVLNFLLYVPELAHSPLYIQDKDGAPVATNAFHSPRWGGIMVYN  
 VDSKTYNASVLPVRVEVDMVRVMEVFLAQLRLLFGIAQPQLPPKCLLSGPTS  
 EGLMTWELDRLLWARSVENLATATTTLTSLAQLLGKISNIVIKDDVASEVYK  
 25 AVAAVQKSAEELASGHLASAFVASQEAVTSSELAFFDPSLLHLLYFPDDQKF  
 AIYIPLFLPMAVPILLSLVKIFLETRKSWRKPEKTD (SEQ ID NO: 165).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in immune and haemopoietic cells, tumors of the ovaries, endometrium, and parathyroid.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune or haemopoietic disorders, particularly cancers, ovarian cancer, cancers of the endocrine system and endometrium, and disorders of the retina. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, haemopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 108 as residues: Phe-11 to Gly-16, Pro-33 to Ser-42. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in T-cells and bone marrow indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders,



such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits

5 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the expression within fetal tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may

10 show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

15       Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have

20 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

25 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the

30 protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2257 of SEQ ID NO:52, b is an integer of 15 to 2271, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 43**

The gene encoding the disclosed cDNA is believed to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 25-44 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

This gene is expressed primarily in infant brain, fetal heart, uterine cancer, colon, metastatic melanoma, spleen, liver, thymus and other cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of developing tissues, haemopoietic or immune system, cardiovascular or musculoskeletal, or neural tissues, uterine cancer and metastatic melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and developing systems, expression of this gene at

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, muscle, immune, hematopoietic, hepatic, developing, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample  
5 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in infant brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of  
10 neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia,  
15 obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system.  
20 Alternatively, expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus  
25 this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy.

The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a  
30 very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g., for treating human

immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g., for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g., for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g., for treating infections, tumors); hemostatic or thrombolytic activity (e.g., for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g., for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. The tissue distribution in uterine cancer and metastatic melanoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of these cancers.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2755 of SEQ ID NO:53, b is an integer of 15 to 2769, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

## 25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 44**

This gene shares sequence homology with Serglycin (see GenSeq accession number Q44278; all references available through this accession are hereby incorporated by reference herein.) Serglycin is involved in the regulation of haematopoietic cell function and development.

30 This gene is expressed primarily in brain frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and cognitive conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system (CNS), expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 110 as residues: Tyr-28 to Cys-40. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of central nervous system disorders, esp. schizophrenia, neurodegenerative and memory disorders. The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to  
5 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
10 related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general  
15 formula of a-b, where a is any integer between 1 to 1375 of SEQ ID NO:54, b is an integer of 15 to 1389, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 45**

20 This gene is expressed primarily in human stomach.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal disorders, particularly gastritis, stomach ulcers, and  
25 stomach cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g.,  
30 gastrointesinal, endothelial, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in stomach cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or  
5 diagnosing disease involving the stomach such as inflammation, ulceration or cancers. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor  
10 marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 734 of SEQ ID NO:55, b is an integer of 15 to 748, where both a and b correspond to the positions of nucleotide  
20 residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 46**

The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in  
25 linkage analysis for chromosome 11.

This gene is expressed primarily in brain and fetal liver, and to a lesser extent, in other cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a  
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, neurological, behavioral, hepatic or immune diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides

are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g.,

5 hepatic, immune, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 112 as residues: Ala-24 to Lys-31. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution of this gene predominantly in fetal liver indicates that the gene could be important for the treatment or detection of immune or

15 hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of neurodegenerative disease states and behavioural

20 disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

25 supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of

30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or



more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 4188 of SEQ ID NO:56, b is an integer of 15 to 4202, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

5

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 47**

Preferred polypeptides of the invention comprise the following amino acid sequence:

KLLLTKEQKLELARLQVDTSGSKEFGTSGIPAKCRFPKIFVNTDDTYEELHLI

10 VYKVTTVFLPAL (SEQ ID NO: 166). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and lymph node of breast cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer and neural disorders, particular neurodegenerative, neurological, or psychological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurological and psychological disorders, including but not limited to: trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, toxic neuropathies induced by neurotoxins, inflammatory diseases such as meningitis and encephalitis, demyelinating diseases, neurodegenerative diseases

such as Parkinson's Disease, Huntington's Disease, Alzheimer's Disease, peripheral neuropathies, multiple sclerosis, neoplasia of neuroectodermal origin, etc.

In addition, the secreted protein can be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to  
5 identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities although no evidence for any is provided in the specification. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g., for treating human  
10 immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of haematopoiesis (e.g., for treating anaemia or as adjunct to chemotherapy); stimulation of growth of bone, cartilage, tendons, ligaments and/or nerves (e.g., for treating wounds, periodontal disease, neurological diseases stroke, fibrosis); inhibition or stimulation of follicle stimulating hormone (for control of  
15 fertility); chemotactic and chemokinetic activities (e.g., for treating infections, tumours); haemostatic or thrombolytic activity (e.g., for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g., for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative disease; for regulation of metabolism, behaviour, and many others. Also contemplated is the  
20 use of the corresponding nucleic acid in gene therapy procedures.

The tissue distribution in cancerous breast tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of breast cancer, and cancer in general. Furthermore, the protein may also be used to determine biological activity, to raise  
25 antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly  
30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 840 of SEQ ID NO:57, b is an integer of 15 to 854, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 48

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLVPQVRTSYNFGRT  
FLGLDKCNACIGTSICKKFFKERNKI (SEQ ID NO: 167). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

QLPLWSPASVQPRVDSQRARGSPPEPKMEPQLGPEAAALRPGWLALLLWVS  
ALSCSFSLPASSLSSLVPQVRTSYNFGRTFLGLDKCNACIGTSICKKFFKKEEIRS  
DNWLASHLGLPPDSLLSYpanySDDSKIWRPVEIFRLVSKYQNEISDRKICAS  
ASAPKTCSIERVLRKTERFQKWLQAKRLTPDLVQDC HQGQRELKFLCMLR  
(SEQ ID NO: 168). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide variant of the invention comprises the following amino acid sequence:

MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLVPQVRTSYNFGRT  
FLGLDKCNACIGTSICKKFFKKEEIRSDNWLASHLGTASRFPLXSYPCKLLQMIX  
KIWXPCGXLLTGQ QXSNEISKQEIXCLLHPPPKNLHIDV (SEQ ID NO: 169).  
Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in endothelial cells, and to a lesser extent, in the adult pulmonary system.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular and pulmonary diseases and/or disorders, particularly atherosclerosis, and microvascular disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

5 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary and vascular systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endothelial, pulmonary, cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, pulmonary surfactant,

10 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic

15 epitopes shown in SEQ ID NO: 114 as residues: Arg-45 to Gly-51, Glu-75 to Asn-81, Ala-99 to Ile-107, Lys-119 to Asp-126, Leu-145 to Gln-152. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution of this gene in the pulmonary system, and in particular endothelial cells, indicates that it could play a role in the treatment/detection of lung

20 lymphoma or sarcoma formation, pulmonary edema and embolism, bronchitis and cystic fibrosis. Expression in endothelial cells suggest a role in the treatment and/or detection of vascular disorders including vasculitis, cardiovascular disorders such as myocardial infarction, myocarditis, ischemia and stroke. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

25 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1441 of SEQ ID NO:58, b is an integer of 15 to 1455, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 49**

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

GPRARVQGFSGADIVKFMALGSMYLVLTIVAKVLRGAEPCCGPLKNRVLRP  
CPLPVHCPLPIPSAEGIPWVAYLPIRWFISCCPGHCIQIPMCTS (SEQ ID NO:

170). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in kidney, and to a lesser extent, in a wide variety of human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal or urogenital diseases and/or disorders, particularly kidney cancer or nephritis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urological or renal system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, urogenital, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Representative uses are described here and elsewhere herein. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 579 of SEQ ID NO:59, b is an integer of 15 to 593, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HKGCN17	203105 08/13/98	pSport1	11	3289	1	3289	218	218	67	1	23	24	155
2	HETAD68	203081 07/30/98	Uni-ZAP XR	12	2342	1	2342	84	84	68	1	32	33	69
3	HPIAT78	203081 07/30/98	Uni-ZAP XR	13	1666	1	1666	30	30	69	1	29	30	501
4	HMWGY65	203105 08/13/98	Uni-ZAP XR	14	2027	1	1976	42	42	70	1	21	22	188
5	HDTAB58	203081 07/30/98	pCMVSPORT 2.0	15	2334	1874	2334	133	133	71	1	33	34	486
5	HDTAB58	203081 07/30/98	pCMVSPORT 2.0	60	496	159	496	223	223	116	1	37	38	48
6	HEOMQ62	203105 08/13/98	pSport1	16	2608	1	2608	130	130	72	1	24	25	87
7	HWLJQ88	203081 07/30/98	pSport1	17	1291	1	1291	114	114	73	1	30	31	299
7	HWLJQ88	203081 07/30/98	pSport1	61	1292	1	1292	403	403	117	1	25	26	203
8	HMICP03	203105 08/13/98	Uni-ZAP XR	18	3129	1	3129	86	86	74	1	34	35	47

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
9	HAJAB01	203105 08/13/98	pCMVSPORT 3.0	19	3629	1	3629	147	147	75	1	15	16	43
10	HE2AT09	203105 08/13/98	Uni-ZAP XR	20	1144	437	1144	435	435	76	1	17	18	140
11	HSDJA15	203081 07/30/98	Uni-ZAP XR	21	1443	1	1443	247	247	77	1	20	21	152
12	HAMGW29	203105 08/13/98	pCMVSPORT 3.0	22	1053	12	1053	25	25	78	1	27	28	179
13	HAPSR85	203081 07/30/98	Uni-ZAP XR	23	741	1	741	428	428	79	1	21	22	69
14	HTOHD42	203081 07/30/98	Uni-ZAP XR	24	946	1	946	155	155	80	1	24	25	190
15	HWLIH65	203081 07/30/98	pSport1	25	831	1	831	129	129	81	1	18	19	165
16	HTOJA73	203105 08/13/98	Uni-ZAP XR	26	1294	1	1294	100	100	82	1	21	22	41
17	HPMGJ45	203105 08/13/98	Uni-ZAP XR	27	1656	1	1656	119	119	83	1	25	26	48
18	HFVIC62	203105 08/13/98	pBluescript	28	1350	1	1350	114	114	84	1	31	32	56
19	HHENW77	203105 08/13/98	pCMVSPORT 3.0	29	1766	1	1714	205	205	85	1	19	20	42



Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
20	HMSIV91	203105 08/13/98	Uni-ZAP XR	30	2790	1	2790	95	95	86	1	19	20	40
21	HMSKC04	203105 08/13/98	Uni-ZAP XR	31	1417	1	1417	133	133	87	1	22	23	73
22	HSAZG33	203105 08/13/98	Uni-ZAP XR	32	1906	1	1906	122	122	88	1	23	24	46
23	HTEBC92	203081 07/30/98	Uni-ZAP XR	33	543	1	543	63	63	89	1	22	23	62
23	HTEBC92	209215 08/21/97	Uni-ZAP XR	62	398	241	398		156	118	1	13	14	18
24	HTXEL29	203081 07/30/98	Uni-ZAP XR	34	1452	1	1452	322	322	90	1	25	26	69
24	HTXEL29	209090 06/05/97	Uni-ZAP XR	63	1202	1	1202		294	119	1			12
25	HDP4W44	203105 08/13/98	pCMVSPORT 3.0	35	2908	1	2908	40	40	91	1	30	31	56
26	HMACS20	203105 08/13/98	Uni-ZAP XR	36	953	1	953	227	227	92	1	43	44	67
27	HAJAY88	203105 08/13/98	pCMVSPORT 3.0	37	3864	1	3864	201	201	93	1	25	26	44
28	HBOEG69	203081 07/30/98	pSPORT1	38	1411	1	1411	302	302	94	1	19	20	54

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
29	HWLEQ37	203081 07/30/98	pSport1	39	1182	1	1182	110	110	95	1	19	20	40
30	HE9CS37	203105 08/13/98	Uni-ZAP XR	40	2457	65	2454	183	183	96	1	22	23	40
31	HNGEI34	203105 08/13/98	Uni-ZAP XR	41	1847	1	1847	70	70	97	1	18	19	49
32	HTOAT76	203105 08/13/98	Uni-ZAP XR	42	2597	1	2597	100	100	98	1	41	42	57
33	HDPVH60	203105 08/13/98	pCMVSPORT 3.0	43	3116	1	3100	8	8	99	1	45	46	51
34	HL YCR65	203105 08/13/98	pSport1	44	3460	1	3460	28	28	100	1	19	20	44
35	HARAY91	203105 08/13/98	pBluescript SK-	45	2622	1	2622	214	214	101	1	46	47	57
36	HCHNT03	203105 08/13/98	pSport1	46	1984	133	1984	228	228	102	1	17	18	40
37	HCUBW95	203105 08/13/98	ZAP Express	47	1987	1	1987	131	131	103	1	20	21	142
38	HDPLV95	203105 08/13/98	pCMVSPORT 3.0	48	2113	1	2113	12	12	104	1	18	19	43
39	HEMGB12	203105 08/13/98	Uni-ZAP XR	49	3465	1	3438	78	78	105	1	48	49	62

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
40	HHENP27	203105 08/13/98	pCMVSPORT 3.0	50	1237	1	1237	12	12	106	1	22	23	282
41	HSPBF70	203105 08/13/98	pSport1	51	1397	288	1397	429	429	107	1	19	20	97
42	HTXKB57	203105 08/13/98	Uni-ZAP XR	52	2271	1	2271	290	290	108	1	16	17	61
43	HUKAA55	203105 08/13/98	Lambda ZAP II	53	2769	129	2769	263	263	109	1	32	33	46
44	HFXGT58	203105 08/13/98	Lambda ZAP II	54	1389	12	1389	238	238	110	1	28	29	46
45	HROAS46	203081 07/30/98	Uni-ZAP XR	55	748	1	748	123	123	111	1	19	20	41
46	HUSFF19	203081 07/30/98	pBluescript	56	4202	863	2447	1080	1080	112	1	20	21	41
47	HBWBX21	203105 08/13/98	ZAP Express	57	854	1	854	201	201	113	1	17	18	43
48	HUVDJ43	203081 07/30/98	Uni-ZAP XR	58	1455	103	1455	128	128	114	1	31	32	182
48	HUVDJ43	203081 07/30/98	Uni-ZAP XR	64	1517	1	1517	133	133	120	1	31	32	79
48	HUVDJ43	203027 06/26/97	Uni-ZAP XR	65	526	69	526	89	89	121	1	31	32	146

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
49	HTLCU49	203081 07/30/98	Uni-ZAP XR	59	593	1	593	170	170	115	1	20	21	80
49	HTLCU49	203027 06/26/97	Uni-ZAP XR	66	664	1	664	249	249	122	1	20	21	80

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X (where X may be any of the polynucleotide sequences disclosed in the sequence listing) and the translated SEQ ID NO:Y (where Y may be any of the polypeptide sequences disclosed in the sequence listing) are sufficiently

accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also

5 hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used, for example, to generate antibodies which bind specifically to proteins containing the polypeptides and the secreted proteins encoded by the cDNA clones identified in Table 1.

10 Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the  
15 actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the  
20 generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods.

25 The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

30 The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed

herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are allelic variants, orthologs, and/or  
5 species homologs. Procedures known in the art can be used to obtain full-length genes, allelic variants, splice variants, full-length coding portions, orthologs, and/or species homologs of genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or a deposited clone, using information from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologs may  
10 be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly  
15 produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).  
20 It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an  
25 isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural, synthetic or  
30 recombinant sources using techniques described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the secreted protein.

The present invention provides a polynucleotide comprising, or alternatively consisting of, the nucleic acid sequence of SEQ ID NO:X, and/or a cDNA contained in ATCC deposit Z. The present invention also provides a polypeptide comprising, or alternatively, consisting of, the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide encoded by the cDNA contained in ATCC deposit Z. Polynucleotides encoding a polypeptide comprising, or alternatively consisting of the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide sequence encoded by the cDNA contained in ATCC deposit Z are also encompassed by the invention.

#### Signal Sequences

The present invention also encompasses mature forms of the polypeptide having the polypeptide sequence of SEQ ID NO:Y and/or the polypeptide sequence encoded by the cDNA in a deposited clone. Polynucleotides encoding the mature forms (such as, for example, the polynucleotide sequence in SEQ ID NO:X and/or the polynucleotide sequence contained in the cDNA of a deposited clone) are also encompassed by the invention. According to the signal hypothesis, proteins secreted by mammalian cells have a signal or secretory leader sequence which is cleaved from the mature protein once export of the growing protein chain across the rough endoplasmic reticulum has been initiated. Most mammalian cells and even insect cells cleave secreted proteins with the same specificity. However, in some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide.

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of



these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. Nonetheless, the present invention provides the mature protein produced by expression of the polynucleotide sequence of SEQ ID NO:X and/or the polynucleotide sequence contained in the cDNA of a deposited clone, in a mammalian cell (e.g., COS cells, as described below). These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

30

#### **Polynucleotide and Polypeptide Variants**

The present invention is directed to variants of the polynucleotide sequence disclosed in SEQ ID NO:X, the complementary strand thereto, and/or the cDNA sequence contained in the deposited clone.

The present invention also encompasses variants of the polypeptide sequence  
5 disclosed in SEQ ID NO:Y and/or encoded by the deposited clone.

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

10 The present invention is also directed to nucleic acid molecules which comprise, or alternatively consist of, a nucleotide sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for example, the nucleotide coding sequence in SEQ ID NO:X or the complementary strand thereto, the nucleotide coding sequence contained in a deposited cDNA clone or the complementary strand  
15 thereto, a nucleotide sequence encoding the polypeptide of SEQ ID NO:Y, a nucleotide sequence encoding the polypeptide encoded by the cDNA contained in a deposited clone, and/or polynucleotide fragments of any of these nucleic acid molecules (e.g., those fragments described herein). Polynucleotides which hybridize to these nucleic acid molecules under stringent hybridization conditions or lower  
20 stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

The present invention is also directed to polypeptides which comprise, or alternatively consist of, an amino acid sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% identical to, for example, the polypeptide sequence  
25 shown in SEQ ID NO:Y, the polypeptide sequence encoded by the cDNA contained in a deposited clone, and/or polypeptide fragments of any of these polypeptides (e.g., those fragments described herein).

By a nucleic acid having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended  
30 that the nucleotide sequence of the nucleic acid is identical to the reference sequence except that the nucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In

other words, to obtain a nucleic acid having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245(1990)). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using

the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the  
5 purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent  
10 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the  
15 deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes  
20 of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid  
25 alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy  
30 terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, an amino acid sequences shown in Table 1 (SEQ ID NO:Y) or to the amino acid sequence encoded by cDNA contained in a deposited clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245(1990)). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10  
5 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100  
10 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not  
15 matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter  
20 the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host  
25 (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These  
30 allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions,

inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies  
5 for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions  
10 where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes  
15 at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

20 As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains  
25 are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr,  
30 and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues,



where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for  
5 example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as, for example, an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of  
10 charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev.  
15 Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid  
20 substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1  
25 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

30

### **Polynucleotide and Polypeptide Fragments**

The present invention is also directed to polynucleotide fragments of the polynucleotides of the invention.

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence which: is a portion of that contained in a deposited clone, or encoding the polypeptide encoded by the cDNA in a deposited clone; is a portion of that shown in SEQ ID NO:X or the complementary strand thereto, or is a portion of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:Y. The nucleotide fragments of the invention are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt, at least about 50 nt, at least about 75 nt, or at least about 150 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in a deposited clone or the nucleotide sequence shown in SEQ ID NO:X. In this context "about" includes the particularly recited value, a value larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. These nucleotide fragments have uses that include, but are not limited to, as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X, or the complementary strand thereto, or the cDNA contained in a deposited clone. In this context "about" includes the particularly recited ranges, and ranges larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein. Polynucleotides which hybridize to these nucleic acid molecules under

stringent hybridization conditions or lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

In the present invention, a "polypeptide fragment" refers to an amino acid sequence which is a portion of that contained in SEQ ID NO:Y or encoded by the cDNA contained in a deposited clone. Protein (polypeptide) fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges or values, and ranges or values larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotides encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are

specifically contemplated by the present invention. Moreover, polynucleotides encoding these domains are also contemplated.

Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity. Polynucleotides encoding these polypeptide fragments are also encompassed by the invention.

#### 10 Epitopes & Antibodies

The present invention is also directed to polypeptide fragments comprising, or alternatively consisting of, an epitope of the polypeptide sequence shown in SEQ ID NO:Y, or the polypeptide sequence encoded by the cDNA contained in a deposited clone. Polynucleotides encoding these epitopes (such as, for example, the sequence disclosed in SEQ ID NO:X) are also encompassed by the invention, is the nucleotide sequences of the complementary strand of the polynucleotides encoding these epitopes. And polynucleotides which hybridize to the complementary strand under stringent hybridization conditions or lower stringency conditions.

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, and most preferably between about 15 to about 30

amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, 5 for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used, for example, to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow et al., Proc. Natl. Acad. Sci. USA 82:910-914; 10 and Bittle et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids 15 have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

Epitope-bearing polypeptides of the present invention may be used to induce antibodies according to methods well known in the art including, but not limited to, *in vivo* immunization, *in vitro* immunization, and phage display methods. See, e.g., 20 Sutcliffe et al., supra; Wilson et al., supra, and Bittle et al., J. Gen. Virol., 66:2347-2354 (1985). If *in vivo* immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to 25 a carrier using a linker such as -maleimidobenzoyl- N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µgs of peptide or carrier 30 protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide

adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

5           As one of skill in the art will appreciate, and discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to heterologous polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, any combination thereof  
10 including both entire domains and portions thereof) resulting in chimeric polypeptides. These fusion proteins facilitate purification, and show an increased half-life *in vivo*. This has been shown, *e.g.*, for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, *e.g.*, EPA  
15 0,394,827; Traunecker et al., *Nature*, 331:84-86 (1988). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion can also be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, *e.g.*, Fountoulakis et al., *J. Biochem.*, 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of  
20 interest as an epitope tag to aid in detection and purification of the expressed polypeptide.

Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to  
25 modulate the activities of polypeptides corresponding to SEQ ID NO:Y thereby effectively generating agonists and antagonists of the polypeptides. See, generally, U.S. Patent Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, and 5,837,458, and Patten, P.A., et al., *Curr. Opin. Biotechnol.* 8:724-33 (1997); Harayama, S., *Trends Biotechnol.* 16(2):76-82 (1998); Hansson, L.O., et al., *J. Mol. Biol.* 287:265-76  
30 (1999); and Lorenzo, M. M. and Blasco, R., *Biotechniques* 24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference). In one embodiment, alteration of polynucleotides corresponding to SEQ ID NO:X and

corresponding polypeptides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments into a desired molecule corresponding to SEQ ID NO:X polynucleotides of the invention by homologous, or site-specific, recombination. In another embodiment, polynucleotides corresponding to SEQ ID NO:X and corresponding polypeptides may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of coding polynucleotide corresponding to SEQ ID NO:X, or the polypeptide encoded thereby may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules.

#### **Antibodies**

The present invention further relates to antibodies and T-cell antigen receptors (TCR) which specifically bind the polypeptides of the present invention. The antibodies of the present invention include IgG (including IgG1, IgG2, IgG3, and IgG4), IgA (including IgA1 and IgA2), IgD, IgE, or IgM, and IgY. As used herein, the term "antibody" (Ab) is meant to include whole antibodies, including single-chain whole antibodies, and antigen-binding fragments thereof. Most preferably the antibodies are human antigen binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')<sub>2</sub>, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a V<sub>L</sub> or V<sub>H</sub> domain. The antibodies may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine, rabbit, goat, guinea pig, camel, horse, or chicken.

Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entire or partial of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are any combinations of variable region(s) and hinge region, CH1, CH2, and CH3 domains. The present invention further includes monoclonal, polyclonal, chimeric, humanized, and human monoclonal and human polyclonal antibodies which specifically bind the polypeptides of the present invention. The present invention

further includes antibodies which are anti-idiotypic to the antibodies of the present invention.

The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for heterologous compositions, such as a heterologous polypeptide or solid support material. *See, e.g.*, WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); US Patents 5,573,920, 4,474,893, 5,601,819, 4,714,681, 4,925,648; Kostelny et al., J. Immunol. 148:1547-1553 (1992).

Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which are recognized or specifically bound by the antibody. The epitope(s) or polypeptide portion(s) may be specified as described herein, *e.g.*, by N-terminal and C-terminal positions, by size in contiguous amino acid residues, or listed in the Tables and Figures. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the exclusion of the same.

Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or homolog of the polypeptides of the present invention are included. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. Further included in the present invention are antibodies which only bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein).

Antibodies of the present invention may also be described or specified in terms of their binding affinity. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-6}M$ ,  $10^{-6}M$ ,  $5 \times 10^{-7}M$ ,  $10^{-7}M$ ,  $5 \times 10^{-8}M$ ,  $10^{-8}M$ ,  $5 \times 10^{-9}M$ ,  $10^{-9}M$ ,  $5 \times 10^{-10}M$ ,  $10^{-10}M$ ,  $5 \times 10^{-11}M$ ,  $10^{-11}M$ ,  $5 \times 10^{-12}M$ ,  $10^{-12}M$ ,  $5 \times 10^{-13}M$ ,  $10^{-13}M$ ,  $5 \times 10^{-14}M$ ,  $10^{-14}M$ ,  $5 \times 10^{-15}M$ ,  $10^{-15}M$ ,  $5 \times 10^{-16}M$ ,  $10^{-16}M$ ,  $5 \times 10^{-17}M$ ,  $10^{-17}M$ ,  $5 \times 10^{-18}M$ ,  $10^{-18}M$ ,  $5 \times 10^{-19}M$ ,  $10^{-19}M$ ,  $5 \times 10^{-20}M$ ,  $10^{-20}M$ ,  $5 \times 10^{-21}M$ ,  $10^{-21}M$ ,  $5 \times 10^{-22}M$ ,  $10^{-22}M$ ,  $5 \times 10^{-23}M$ ,  $10^{-23}M$ ,  $5 \times 10^{-24}M$ ,  $10^{-24}M$ ,  $5 \times 10^{-25}M$ ,  $10^{-25}M$ ,  $5 \times 10^{-26}M$ ,  $10^{-26}M$ ,  $5 \times 10^{-27}M$ ,  $10^{-27}M$ ,  $5 \times 10^{-28}M$ ,  $10^{-28}M$ ,  $5 \times 10^{-29}M$ ,  $10^{-29}M$ ,  $5 \times 10^{-30}M$ ,  $10^{-30}M$ ,  $5 \times 10^{-31}M$ ,  $10^{-31}M$ ,  $5 \times 10^{-32}M$ ,  $10^{-32}M$ ,  $5 \times 10^{-33}M$ ,  $10^{-33}M$ ,  $5 \times 10^{-34}M$ ,  $10^{-34}M$ ,  $5 \times 10^{-35}M$ ,  $10^{-35}M$ ,  $5 \times 10^{-36}M$ ,  $10^{-36}M$ ,  $5 \times 10^{-37}M$ ,  $10^{-37}M$ ,  $5 \times 10^{-38}M$ ,  $10^{-38}M$ ,  $5 \times 10^{-39}M$ ,  $10^{-39}M$ ,  $5 \times 10^{-40}M$ ,  $10^{-40}M$ ,  $5 \times 10^{-41}M$ ,  $10^{-41}M$ ,  $5 \times 10^{-42}M$ ,  $10^{-42}M$ ,  $5 \times 10^{-43}M$ ,  $10^{-43}M$ ,  $5 \times 10^{-44}M$ ,  $10^{-44}M$ ,  $5 \times 10^{-45}M$ ,  $10^{-45}M$ ,  $5 \times 10^{-46}M$ ,  $10^{-46}M$ ,  $5 \times 10^{-47}M$ ,  $10^{-47}M$ ,  $5 \times 10^{-48}M$ ,  $10^{-48}M$ ,  $5 \times 10^{-49}M$ ,  $10^{-49}M$ ,  $5 \times 10^{-50}M$ ,  $10^{-50}M$ ,  $5 \times 10^{-51}M$ ,  $10^{-51}M$ ,  $5 \times 10^{-52}M$ ,  $10^{-52}M$ ,  $5 \times 10^{-53}M$ ,  $10^{-53}M$ ,  $5 \times 10^{-54}M$ ,  $10^{-54}M$ ,  $5 \times 10^{-55}M$ ,  $10^{-55}M$ ,  $5 \times 10^{-56}M$ ,  $10^{-56}M$ ,  $5 \times 10^{-57}M$ ,  $10^{-57}M$ ,  $5 \times 10^{-58}M$ ,  $10^{-58}M$ ,  $5 \times 10^{-59}M$ ,  $10^{-59}M$ ,  $5 \times 10^{-60}M$ ,  $10^{-60}M$ ,  $5 \times 10^{-61}M$ ,  $10^{-61}M$ ,  $5 \times 10^{-62}M$ ,  $10^{-62}M$ ,  $5 \times 10^{-63}M$ ,  $10^{-63}M$ ,  $5 \times 10^{-64}M$ ,  $10^{-64}M$ ,  $5 \times 10^{-65}M$ ,  $10^{-65}M$ ,  $5 \times 10^{-66}M$ ,  $10^{-66}M$ ,  $5 \times 10^{-67}M$ ,  $10^{-67}M$ ,  $5 \times 10^{-68}M$ ,  $10^{-68}M$ ,  $5 \times 10^{-69}M$ ,  $10^{-69}M$ ,  $5 \times 10^{-70}M$ ,  $10^{-70}M$ ,  $5 \times 10^{-71}M$ ,  $10^{-71}M$ ,  $5 \times 10^{-72}M$ ,  $10^{-72}M$ ,  $5 \times 10^{-73}M$ ,  $10^{-73}M$ ,  $5 \times 10^{-74}M$ ,  $10^{-74}M$ ,  $5 \times 10^{-75}M$ ,  $10^{-75}M$ ,  $5 \times 10^{-76}M$ ,  $10^{-76}M$ ,  $5 \times 10^{-77}M$ ,  $10^{-77}M$ ,  $5 \times 10^{-78}M$ ,  $10^{-78}M$ ,  $5 \times 10^{-79}M$ ,  $10^{-79}M$ ,  $5 \times 10^{-80}M$ ,  $10^{-80}M$ ,  $5 \times 10^{-81}M$ ,  $10^{-81}M$ ,  $5 \times 10^{-82}M$ ,  $10^{-82}M$ ,  $5 \times 10^{-83}M$ ,  $10^{-83}M$ ,  $5 \times 10^{-84}M$ ,  $10^{-84}M$ ,  $5 \times 10^{-85}M$ ,  $10^{-85}M$ ,  $5 \times 10^{-86}M$ ,  $10^{-86}M$ ,  $5 \times 10^{-87}M$ ,  $10^{-87}M$ ,  $5 \times 10^{-88}M$ ,  $10^{-88}M$ ,  $5 \times 10^{-89}M$ ,  $10^{-89}M$ ,  $5 \times 10^{-90}M$ ,  $10^{-90}M$ ,  $5 \times 10^{-91}M$ ,  $10^{-91}M$ ,  $5 \times 10^{-92}M$ ,  $10^{-92}M$ ,  $5 \times 10^{-93}M$ ,  $10^{-93}M$ ,  $5 \times 10^{-94}M$ ,  $10^{-94}M$ ,  $5 \times 10^{-95}M$ ,  $10^{-95}M$ ,  $5 \times 10^{-96}M$ ,  $10^{-96}M$ ,  $5 \times 10^{-97}M$ ,  $10^{-97}M$ ,  $5 \times 10^{-98}M$ ,  $10^{-98}M$ ,  $5 \times 10^{-99}M$ ,  $10^{-99}M$ ,  $5 \times 10^{-100}M$ ,  $10^{-100}M$ ,  $5 \times 10^{-101}M$ ,  $10^{-101}M$ ,  $5 \times 10^{-102}M$ ,  $10^{-102}M$ ,  $5 \times 10^{-103}M$ ,  $10^{-103}M$ ,  $5 \times 10^{-104}M$ ,  $10^{-104}M$ ,  $5 \times 10^{-105}M$ ,  $10^{-105}M$ ,  $5 \times 10^{-106}M$ ,  $10^{-106}M$ ,  $5 \times 10^{-107}M$ ,  $10^{-107}M$ ,  $5 \times 10^{-108}M$ ,  $10^{-108}M$ ,  $5 \times 10^{-109}M$ ,  $10^{-109}M$ ,  $5 \times 10^{-110}M$ ,  $10^{-110}M$ ,  $5 \times 10^{-111}M$ ,  $10^{-111}M$ ,  $5 \times 10^{-112}M$ ,  $10^{-112}M$ ,  $5 \times 10^{-113}M$ ,  $10^{-113}M$ ,  $5 \times 10^{-114}M$ ,  $10^{-114}M$ ,  $5 \times 10^{-115}M$ ,  $10^{-115}M$ ,  $5 \times 10^{-116}M$ ,  $10^{-116}M$ ,  $5 \times 10^{-117}M$ ,  $10^{-117}M$ ,  $5 \times 10^{-118}M$ ,  $10^{-118}M$ ,  $5 \times 10^{-119}M$ ,  $10^{-119}M$ ,  $5 \times 10^{-120}M$ ,  $10^{-120}M$ ,  $5 \times 10^{-121}M$ ,  $10^{-121}M$ ,  $5 \times 10^{-122}M$ ,  $10^{-122}M$ ,  $5 \times 10^{-123}M$ ,  $10^{-123}M$ ,  $5 \times 10^{-124}M$ ,  $10^{-124}M$ ,  $5 \times 10^{-125}M$ ,  $10^{-125}M$ ,  $5 \times 10^{-126}M$ ,  $10^{-126}M$ ,  $5 \times 10^{-127}M$ ,  $10^{-127}M$ ,  $5 \times 10^{-128}M$ ,  $10^{-128}M$ ,  $5 \times 10^{-129}M$ ,  $10^{-129}M$ ,  $5 \times 10^{-130}M$ ,  $10^{-130}M$ ,  $5 \times 10^{-131}M$ ,  $10^{-131}M$ ,  $5 \times 10^{-132}M$ ,  $10^{-132}M$ ,  $5 \times 10^{-133}M$ ,  $10^{-133}M$ ,  $5 \times 10^{-134}M$ ,  $10^{-134}M$ ,  $5 \times 10^{-135}M$ ,  $10^{-135}M$ ,  $5 \times 10^{-136}M$ ,  $10^{-136}M$ ,  $5 \times 10^{-137}M$ ,  $10^{-137}M$ ,  $5 \times 10^{-138}M$ ,  $10^{-138}M$ ,  $5 \times 10^{-139}M$ ,  $10^{-139}M$ ,  $5 \times 10^{-140}M$ ,  $10^{-140}M$ ,  $5 \times 10^{-141}M$ ,  $10^{-141}M$ ,  $5 \times 10^{-142}M$ ,  $10^{-142}M$ ,  $5 \times 10^{-143}M$ ,  $10^{-143}M$ ,  $5 \times 10^{-144}M$ ,  $10^{-144}M$ ,  $5 \times 10^{-145}M$ ,  $10^{-145}M$ ,  $5 \times 10^{-146}M$ ,  $10^{-146}M$ ,  $5 \times 10^{-147}M$ ,  $10^{-147}M$ ,  $5 \times 10^{-148}M$ ,  $10^{-148}M$ ,  $5 \times 10^{-149}M$ ,  $10^{-149}M$ ,  $5 \times 10^{-150}M$ ,  $10^{-150}M$ ,  $5 \times 10^{-151}M$ ,  $10^{-151}M$ ,  $5 \times 10^{-152}M$ ,  $10^{-152}M$ ,  $5 \times 10^{-153}M$ ,  $10^{-153}M$ ,  $5 \times 10^{-154}M$ ,  $10^{-154}M$ ,  $5 \times 10^{-155}M$ ,  $10^{-155}M$ ,  $5 \times 10^{-156}M$ ,  $10^{-156}M$ ,  $5 \times 10^{-157}M$ ,  $10^{-157}M$ ,  $5 \times 10^{-158}M$ ,  $10^{-158}M$ ,  $5 \times 10^{-159}M$ ,  $10^{-159}M$ ,  $5 \times 10^{-160}M$ ,  $10^{-160}M$ ,  $5 \times 10^{-161}M$ ,  $10^{-161}M$ ,  $5 \times 10^{-162}M$ ,  $10^{-162}M$ ,  $5 \times 10^{-163}M$ ,  $10^{-163}M$ ,  $5 \times 10^{-164}M$ ,  $10^{-164}M$ ,  $5 \times 10^{-165}M$ ,  $10^{-165}M$ ,  $5 \times 10^{-166}M$ ,  $10^{-166}M$ ,  $5 \times 10^{-167}M$ ,  $10^{-167}M$ ,  $5 \times 10^{-168}M$ ,  $10^{-168}M$ ,  $5 \times 10^{-169}M$ ,  $10^{-169}M$ ,  $5 \times 10^{-170}M$ ,  $10^{-170}M$ ,  $5 \times 10^{-171}M$ ,  $10^{-171}M$ ,  $5 \times 10^{-172}M$ ,  $10^{-172}M$ ,  $5 \times 10^{-173}M$ ,  $10^{-173}M$ ,  $5 \times 10^{-174}M$ ,  $10^{-174}M$ ,  $5 \times 10^{-175}M$ ,  $10^{-175}M$ ,  $5 \times 10^{-176}M$ ,  $10^{-176}M$ ,  $5 \times 10^{-177}M$ ,  $10^{-177}M$ ,  $5 \times 10^{-178}M$ ,  $10^{-178}M$ ,  $5 \times 10^{-179}M$ ,  $10^{-179}M$ ,  $5 \times 10^{-180}M$ ,  $10^{-180}M$ ,  $5 \times 10^{-181}M$ ,  $10^{-181}M$ ,  $5 \times 10^{-182}M$ ,  $10^{-182}M$ ,  $5 \times 10^{-183}M$ ,  $10^{-183}M$ ,  $5 \times 10^{-184}M$ ,  $10^{-184}M$ ,  $5 \times 10^{-185}M$ ,  $10^{-185}M$ ,  $5 \times 10^{-186}M$ ,  $10^{-186}M$ ,  $5 \times 10^{-187}M$ ,  $10^{-187}M$ ,  $5 \times 10^{-188}M$ ,  $10^{-188}M$ ,  $5 \times 10^{-189}M$ ,  $10^{-189}M$ ,  $5 \times 10^{-190}M$ ,  $10^{-190}M$ ,  $5 \times 10^{-191}M$ ,  $10^{-191}M$ ,  $5 \times 10^{-192}M$ ,  $10^{-192}M$ ,  $5 \times 10^{-193}M$ ,  $10^{-193}M$ ,  $5 \times 10^{-194}M$ ,  $10^{-194}M$ ,  $5 \times 10^{-195}M$ ,  $10^{-195}M$ ,  $5 \times 10^{-196}M$ ,  $10^{-196}M$ ,  $5 \times 10^{-197}M$ ,  $10^{-197}M$ ,  $5 \times 10^{-198}M$ ,  $10^{-198}M$ ,  $5 \times 10^{-199}M$ ,  $10^{-199}M$ ,  $5 \times 10^{-200}M$ ,  $10^{-200}M$ ,  $5 \times 10^{-201}M$ ,  $10^{-201}M$ ,  $5 \times 10^{-202}M$ ,  $10^{-202}M$ ,  $5 \times 10^{-203}M$ ,  $10^{-203}M$ ,  $5 \times 10^{-204}M$ ,  $10^{-204}M$ ,  $5 \times 10^{-205}M$ ,  $10^{-205}M$ ,  $5 \times 10^{-206}M$ ,  $10^{-206}M$ ,  $5 \times 10^{-207}M$ ,  $10^{-207}M$ ,  $5 \times 10^{-208}M$ ,  $10^{-208}M$ ,  $5 \times 10^{-209}M$ ,  $10^{-209}M$ ,  $5 \times 10^{-210}M$ ,  $10^{-210}M$ ,  $5 \times 10^{-211}M$ ,  $10^{-211}M$ ,  $5 \times 10^{-212}M$ ,  $10^{-212}M$ ,  $5 \times 10^{-213}M$ ,  $10^{-213}M$ ,  $5 \times 10^{-214}M$ ,  $10^{-214}M$ ,  $5 \times 10^{-215}M$ ,  $10^{-215}M$ ,  $5 \times 10^{-216}M$ ,  $10^{-216}M$ ,  $5 \times 10^{-217}M$ ,  $10^{-217}M$ ,  $5 \times 10^{-218}M$ ,  $10^{-218}M$ ,  $5 \times 10^{-219}M$ ,  $10^{-219}M$ ,  $5 \times 10^{-220}M$ ,  $10^{-220}M$ ,  $5 \times 10^{-221}M$ ,  $10^{-221}M$ ,  $5 \times 10^{-222}M$ ,  $10^{-222}M$ ,  $5 \times 10^{-223}M$ ,  $10^{-223}M$ ,  $5 \times 10^{-224}M$ ,  $10^{-224}M$ ,  $5 \times 10^{-225}M$ ,  $10^{-225}M$ ,  $5 \times 10^{-226}M$ ,  $10^{-226}M$ ,  $5 \times 10^{-227}M$ ,  $10^{-227}M$ ,  $5 \times 10^{-228}M$ ,  $10^{-228}M$ ,  $5 \times 10^{-229}M$ ,  $10^{-229}M$ ,  $5 \times 10^{-230}M$ ,  $10^{-230}M$ ,  $5 \times 10^{-231}M$ ,  $10^{-231}M$ ,  $5 \times 10^{-232}M$ ,  $10^{-232}M$ ,  $5 \times 10^{-233}M$ ,  $10^{-233}M$ ,  $5 \times 10^{-234}M$ ,  $10^{-234}M$ ,  $5 \times 10^{-235}M$ ,  $10^{-235}M$ ,  $5 \times 10^{-236}M$ ,  $10^{-236}M$ ,  $5 \times 10^{-237}M$ ,  $10^{-237}M$ ,  $5 \times 10^{-238}M$ ,  $10^{-238}M$ ,  $5 \times 10^{-239}M$ ,  $10^{-239}M$ ,  $5 \times 10^{-240}M$ ,  $10^{-240}M$ ,  $5 \times 10^{-241}M$ ,  $10^{-241}M$ ,  $5 \times 10^{-242}M$ ,  $10^{-242}M$ ,  $5 \times 10^{-243}M$ ,  $10^{-243}M$ ,  $5 \times 10^{-244}M$ ,  $10^{-244}M$ ,  $5 \times 10^{-245}M$ ,  $10^{-245}M$ ,  $5 \times 10^{-246}M$ ,  $10^{-246}M$ ,  $5 \times 10^{-247}M$ ,  $10^{-247}M$ ,  $5 \times 10^{-248}M$ ,  $10^{-248}M$ ,  $5 \times 10^{-249}M$ ,  $10^{-249}M$ ,  $5 \times 10^{-250}M$ ,  $10^{-250}M$ ,  $5 \times 10^{-251}M$ ,  $10^{-251}M$ ,  $5 \times 10^{-252}M$ ,  $10^{-252}M$ ,  $5 \times 10^{-253}M$ ,  $10^{-253}M$ ,  $5 \times 10^{-254}M$ ,  $10^{-254}M$ ,  $5 \times 10^{-255}M$ ,  $10^{-255}M$ ,  $5 \times 10^{-256}M$ ,  $10^{-256}M$ ,  $5 \times 10^{-257}M$ ,  $10^{-257}M$ ,  $5 \times 10^{-258}M$ ,  $10^{-258}M$ ,  $5 \times 10^{-259}M$ ,  $10^{-259}M$ ,  $5 \times 10^{-260}M$ ,  $10^{-260}M$ ,  $5 \times 10^{-261}M$ ,  $10^{-261}M$ ,  $5 \times 10^{-262}M$ ,  $10^{-262}M$ ,  $5 \times 10^{-263}M$ ,  $10^{-263}M$ ,  $5 \times 10^{-264}M$ ,  $10^{-264}M$ ,  $5 \times 10^{-265}M$ ,  $10^{-265}M$ ,  $5 \times 10^{-266}M$ ,  $10^{-266}M$ ,  $5 \times 10^{-267}M$ ,  $10^{-267}M$ ,  $5 \times 10^{-268}M$ ,  $10^{-268}M$ ,  $5 \times 10^{-269}M$ ,  $10^{-269}M$ ,  $5 \times 10^{-270}M$ ,  $10^{-270}M$ ,  $5 \times 10^{-271}M$ ,  $10^{-271}M$ ,  $5 \times 10^{-272}M$ ,  $10^{-272}M$ ,  $5 \times 10^{-273}M$ ,  $10^{-273}M$ ,  $5 \times 10^{-274}M$ ,  $10^{-274}M$ ,  $5 \times 10^{-275}M$ ,  $10^{-275}M$ ,  $5 \times 10^{-276}M$ ,  $10^{-276}M$ ,  $5 \times 10^{-277}M$ ,  $10^{-277}M$ ,  $5 \times 10^{-278}M$ ,  $10^{-278}M$ ,  $5 \times 10^{-279}M$ ,  $10^{-279}M$ ,  $5 \times 10^{-280}M$ ,  $10^{-280}M$ ,  $5 \times 10^{-281}M$ ,  $10^{-281}M$ ,  $5 \times 10^{-282}M$ ,  $10^{-282}M$ ,  $5 \times 10^{-283}M$ ,  $10^{-283}M$ ,  $5 \times 10^{-284}M$ ,  $10^{-284}M$ ,  $5 \times 10^{-285}M$ ,  $10^{-285}M$ ,  $5 \times 10^{-286}M$ ,  $10^{-286}M$ ,  $5 \times 10^{-287}M$ ,  $10^{-287}M$ ,  $5 \times 10^{-288}M$ ,  $10^{-288}M$ ,  $5 \times 10^{-289}M$ ,  $10^{-289}M$ ,  $5 \times 10^{-290}M$ ,  $10^{-290}M$ ,  $5 \times 10^{-291}M$ ,  $10^{-291}M$ ,  $5 \times 10^{-292}M$ ,  $10^{-292}M$ ,  $5 \times 10^{-293}M$ ,  $10^{-293}M$ ,  $5 \times 10^{-294}M$ ,  $10^{-294}M$ ,  $5 \times 10^{-295}M$ ,  $10^{-295}M$ ,  $5 \times 10^{-296}M$ ,  $10^{-296}M$ ,  $5 \times 10^{-297}M$ ,  $10^{-297}M$ ,  $5 \times 10^{-298}M$ ,  $10^{-298}M$ ,  $5 \times 10^{-299}M$ ,  $10^{-299}M$ ,  $5 \times 10^{-300}M$ ,  $10^{-300}M$ ,  $5 \times 10^{-301}M$ ,  $10^{-301}M$ ,  $5 \times 10^{-302}M$ ,  $10^{-302}M$ ,  $5 \times 10^{-303}M$ ,  $10^{-303}M$ ,  $5 \times 10^{-304}M$ ,  $10^{-304}M$ ,  $5 \times 10^{-305}M$ ,  $10^{-305}M$ ,  $5 \times 10^{-306}M$ ,  $10^{-306}M$ ,  $5 \times 10^{-307}M$ ,  $10^{-307}M$ ,  $5 \times 10^{-308}M$ ,  $10^{-308}M$ ,  $5 \times 10^{-309}M$ ,  $10^{-309}M$ ,  $5 \times 10^{-310}M$ ,  $10^{-310}M$ ,  $5 \times 10^{-311}M$ ,  $10^{-311}M$ ,  $5 \times 10^{-312}M$ ,  $10^{-312}M$ ,  $5 \times 10^{-313}M$ ,  $10^{-313}M$ ,  $5 \times 10^{-314}M$ ,  $10^{-314}M$ ,  $5 \times 10^{-315}M$ ,  $10^{-315}M$ ,  $5 \times 10^{-316}M$ ,  $10^{-316}M$ ,  $5 \times 10^{-317}M$ ,  $10^{-317}M$ ,  $5 \times 10^{-318}M$ ,  $10^{-318}M$ ,  $5 \times 10^{-319}M$ ,  $10^{-319}M$ ,  $5 \times 10^{-320}M$ ,  $10^{-320}M$ ,  $5 \times 10^{-321}M$ ,  $10^{-321}M$ ,  $5 \times 10^{-322}M$ ,  $10^{-322}M$ ,  $5 \times 10^{-323}M$ ,  $10^{-323}M$ ,  $5 \times 10^{-324}M$ ,  $10^{-324}M$ ,  $5 \times 10^{-325}M$ ,  $10^{-325}M$ ,  $5 \times 10^{-326}M$ ,  $10^{-326}M$ ,  $5 \times 10^{-327}M$ ,  $10^{-327}M$ ,  $5 \times 10^{-328}M$ ,  $10^{-328}M$ ,  $5 \times 10^{-329}M$ ,  $10^{-329}M$ ,  $5 \times 10^{-330}M$ ,  $10^{-330}M$ ,  $5 \times 10^{-331}M$ ,  $10^{-331}M$ ,  $5 \times 10^{-332}M$ ,  $10^{-332}M$ ,  $5 \times 10^{-333}M$ ,  $10^{-333}M$ ,  $5 \times 10^{-334}M$ ,  $10^{-334}M$ ,  $5 \times 10^{-335}M$ ,  $10^{-335}M$ ,  $5 \times 10^{-336}M$ ,  $10^{-336}M$ ,  $5 \times 10^{-337}M$ ,  $10^{-337}M$ ,  $5 \times 10^{-338}M$ ,  $10^{-338}M$ ,  $5 \times 10^{-339}M$ ,  $10^{-339}M$ ,  $5 \times 10^{-340}M$ ,  $10^{-340}M$ ,  $5 \times 10^{-341}M$ ,  $10^{-341}M$ ,  $5 \times 10^{-342}M$ ,  $10^{-342}M$ ,  $5 \times 10^{-343}M$ ,  $10^{-343}M$ ,  $5 \times 10^{-344}M$ ,  $10^{-344}M$ ,  $5 \times 10^{-345}M$ ,  $10^{-345}M$ ,  $5 \times 10^{-346}M$ ,  $10^{-346}M$ ,  $5 \times 10^{-347}M$ ,  $10^{-347}M$ ,  $5 \times 10^{-348}M$ ,  $10^{-348}M$ ,  $5 \times 10^{-349}M$ ,  $10^{-349}M$ ,  $5 \times 10^{-350}M$ ,  $10^{-350}M$ ,  $5 \times 10^{-351}M$ ,  $10^{-351}M$ ,  $5 \times 10^{-352}M$ ,  $10^{-352}M$ ,  $5 \times 10^{-353}M$ ,  $10^{-353}M$ ,  $5 \times 10^{-354}M$ ,  $10^{-354}M$ ,  $5 \times 10^{-355}M$ ,  $10^{-355}M$ ,  $5 \times 10^{-356}M$ ,  $10^{-356}M$ ,  $5 \times 10^{-357}M$ ,  $10^{-357}M$ ,  $5 \times 10^{-358}M$ ,  $10^{-358}M$ ,  $5 \times 10^{-359}M$ ,  $10^{-359}M$ ,  $5 \times 10^{-360}M$ ,  $10^{-360}M$ ,  $5 \times 10^{-361}M$ ,  $10^{-361}M$ ,  $5 \times 10^{-362}M$ ,  $10^{-362}M$ ,  $5 \times 10^{-363}M$ ,  $10^{-363}M$ ,  $5 \times 10^{-364}M$ ,  $10^{-364}M$ ,  $5 \times 10^{-365}M$ ,  $10^{-365}M$ ,  $5 \times 10^{-366}M$ ,  $10^{-366}M$ ,  $5 \times 10^{-367}M$ ,  $10^{-367}M$ ,  $5 \times 10^{-368}M$ ,  $10^{-368}M$ ,  $5 \times 10^{-369}M$ ,  $10^{-369}M$ ,  $5 \times 10^{-370}M$ ,



$^9\text{M}$ ,  $10^{-9}\text{M}$ ,  $5\times 10^{-10}\text{M}$ ,  $10^{-10}\text{M}$ ,  $5\times 10^{-11}\text{M}$ ,  $10^{-11}\text{M}$ ,  $5\times 10^{-12}\text{M}$ ,  $10^{-12}\text{M}$ ,  $5\times 10^{-13}\text{M}$ ,  $10^{-13}\text{M}$ ,  $5\times 10^{-14}\text{M}$ ,  $10^{-14}\text{M}$ ,  $5\times 10^{-15}\text{M}$ , and  $10^{-15}\text{M}$ .

Antibodies of the present invention have uses that include, but are not limited to, methods known in the art to purify, detect, and target the polypeptides of the present invention including both *in vitro* and *in vivo* diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference in the entirety).

The antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, or toxins. See, e.g., WO 92/08495; WO 91/14438; WO 89/12624; US Patent 5,314,995; and EP 0 396 387.

The antibodies of the present invention may be prepared by any suitable method known in the art. For example, a polypeptide of the present invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. The term "monoclonal antibody" is not a limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced. Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technology.

Hybridoma techniques include those known in the art and taught in Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: MONOCLONAL ANTIBODIES AND

T-CELL HYBRIDOMAS 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). Fab and F(ab')<sub>2</sub> fragments may be produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments).

5           Alternatively, antibodies of the present invention can be produced through the application of recombinant DNA and phage display technology or through synthetic chemistry using methods known in the art. For example, the antibodies of the present invention can be prepared using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of a  
10   phage particle which carries polynucleotide sequences encoding them. Phage with a desired binding property are selected from a repertoire or combinatorial antibody library (e.g. human or murine) by selecting directly with antigen, typically antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 with Fab, Fv or disulfide stabilized  
15   Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-  
20   18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT/GB91/01134; WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and US Patents 5,698,426, 5,223,409, 5,403,484, 5,580,717, 5,427,908, 5,750,753, 5,821,047, 5,571,698, 5,427,908, 5,516,637, 5,780,225, 5,658,727 and 5,733,743 (said references incorporated by  
25   reference in their entireties).

          As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host including mammalian cells, insect cells, plant cells,  
30   yeast, and bacteria. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sub>2</sub> fragments can also be employed using methods known in the art such as those disclosed in WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and

Sawai et al., AJRI 34:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu, L. et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988). For some uses, including *in vivo* use of antibodies in humans and *in vitro* detection assays, it may be preferable to use chimeric, humanized, or human antibodies. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Gillies et al., (1989) J. Immunol. Methods 125:191-202; and US Patent 5,807,715. Antibodies can be humanized using a variety of techniques including CDR-grafting (EP 0 239 400; WO 91/09967; US Patent 5,530,101; and 5,585,089), veneering or resurfacing (EP 0 592 106; EP 0 519 596; Padlan E.A., Molecular Immunology 28(4/5):489-498 (1991); Studnicka et al., Protein Engineering 7(6):805-814 (1994); Roguska. et al., PNAS 91:969-973 (1994)), and chain shuffling (US Patent 5,565,332). Human antibodies can be made by a variety of methods known in the art including phage display methods described above. See also, US Patents 4,444,887, 4,716,111, 5,545,806, and 5,814,318; and WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741 (said references incorporated by reference in their entireties).

Further included in the present invention are antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide of the present invention. The antibodies may be specific for antigens other than polypeptides of the present invention. For example, antibodies may be used to target the polypeptides of the present invention to particular cell types, either *in vitro* or *in vivo*, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in *in vitro* immunoassays and purification methods using methods known in the art. See e.g., Harbor et al. *supra* and WO 93/21232; EP 0 439 095; Naramura et al., Immunol. Lett. 39:91-99 (1994); US Patent 5,474,981; Gillies et al., PNAS 89:1428-1432 (1992);

Fell et al., J. Immunol. 146:2446-2452 (1991) (said references incorporated by reference in their entireties).

The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains or portions thereof. The polypeptides of the present invention may be fused or conjugated to the above antibody portions to increase the *in vivo* half life of the polypeptides or for use in immunoassays using methods known in the art. The polypeptides may also be fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. See e.g., US Patents 5,336,603, 5,622,929, 5,359,046, 5,349,053, 5,447,851, 5,112,946; EP 0 307 434, EP 0 367 166; WO 96/04388, WO 91/06570; Ashkenazi et al., PNAS 88:10535-10539 (1991); Zheng et al., J. Immunol. 154:5590-5600 (1995); and Vil et al., PNAS 89:11337-11341 (1992) (said references incorporated by reference in their entireties).

The invention further relates to antibodies which act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Included are both receptor-specific antibodies and ligand-specific antibodies. Included are receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. Also include are receptor-specific antibodies which both prevent ligand binding and receptor activation. Likewise, included are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor

activation, but do not prevent the ligand from binding the receptor. Further included are antibodies which activate the receptor. These antibodies may act as agonists for either all or less than all of the biological activities affected by ligand-mediated receptor activation. The antibodies may be specified as agonists or antagonists for biological activities comprising specific activities disclosed herein. The above antibody agonists can be made using methods known in the art. See e.g., WO 96/40281; US Patent 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen, et al., Cancer Res. 58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon, et al., J. Immunol. 160(7):3170-3179 (1998); Prat et al., J. Cell. Sci. 111(Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997); Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al., Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998); Bartunek et al., Cytokine 8(1):14-20 (1996) (said references incorporated by reference in their entireties).

As discussed above, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, FASEB J. 7(5):437-444; (1989) and Nissinoff, J. Immunol. 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

### **Fusion Proteins**

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a

second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for

example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for  
5 example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker  
10 sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for  
15 instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides  
20 or the polypeptides of the present invention.

#### **Vectors, Host Cells, and Protein Production**

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant  
25 techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker  
30 for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the

vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, phoA and tac  
5 promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a  
10 translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin  
15 resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture  
20 mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech,  
25 Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated  
30 transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that the



polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or  
5 ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

10 Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant,  
15 insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine  
20 encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

25 In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide  
30 sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control

regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al.,  
5 Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, Proteins: Structures and  
10 Molecular Principles, W.H. Freeman & Co., N.Y., and Hunkapiller et al., Nature, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide sequence of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid  
15 Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-  
20 butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoro-amino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

The invention encompasses polypeptides which are differentially modified  
25 during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin,  
30 papain, V8 protease, NaBH<sub>4</sub>; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and  
5 addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Also provided by the invention are chemically modified derivatives of the  
10 polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent NO: 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose,  
15 dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between  
20 about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or  
25 lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those  
30 skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may

be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues;  
5 those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

10 One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be  
15 performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be  
20 accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

25 The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, pharmaceutical compositions) containing them. In specific embodiments, the polypeptides of the invention are monomers,  
30 dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

Multimers encompassed by the invention may be homomers or heteromers.

As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of SEQ ID NO:Y or encoded by the cDNA contained in a deposited clone (including fragments, variants, splice variants, and fusion proteins, corresponding to these polypeptides as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (*e.g.*, containing polypeptides having identical or different amino acid sequences) or a homotrimer (*e.g.*, containing polypeptides having identical and/or different amino acid sequences). In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (*i.e.*, polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid

residues contained in the polypeptide sequence ( e.g., that recited in the sequence listing, or contained in the polypeptide encoded by a deposited clone). In one instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally  
5 occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the heterologous polypeptide sequence in a fusion protein of the invention.

In one example, covalent associations are between the heterologous sequence  
10 contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in an Fc fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from  
15 another protein that is capable of forming covalently associated multimers, such as for example, osteopontin (see, e.g., International Publication NO: WO 98/49305, the contents of which are herein incorporated by reference in its entirety). In another embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627  
20 (hereby incorporated by reference). Proteins comprising multiple polypeptides of the invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper  
25 polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and  
30 derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant

fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using techniques known in the art.

5 Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by  
10 reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide sequence. In a further embodiment, associations  
15 proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers  
20 of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, multimers of the invention may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues  
25 located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers  
30 containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the

polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Alternatively, multimers of the invention may be generated using genetic  
5 engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are  
10 generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is herein incorporated by  
15 reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is herein incorporated by  
20 reference in its entirety).

#### Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes  
25 known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present  
30 invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be



selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an  
5 amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping  
10 strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread.  
15 This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to  
20 mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

25 Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins  
30 University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural  
5 alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this  
10 polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic  
15 marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to  
20 either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991) ) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally  
25 results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One  
30 goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate

manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of  
5 restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification  
10 difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for  
15 amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

20 Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic  
25 biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for  
30 forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with

tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures  
5 for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making  
10 oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

#### Uses of the Polypeptides

15 Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M.,  
20 et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such  
25 as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and technetium ( $^{99\text{m}}\text{Tc}$ ), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo  
30 imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers

for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with  
5 an appropriate detectable imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99\text{m}}\text{Tc}$ ), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety  
10 needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{m}}\text{Tc}$ . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics  
15 of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells  
20 or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease.  
25 For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the  
30 activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

#### Gene Therapy Methods

Another aspect of the present invention is to gene therapy methods for treating disorders, diseases and conditions. The gene therapy methods relate to the introduction of nucleic acid (DNA, RNA and antisense DNA or RNA) sequences into an animal to achieve expression of a polypeptide of the present invention. This method requires a polynucleotide which codes for a polypeptide of the invention that operatively linked to a promoter and any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques are known in the art, see, for example, WO90/11092, which is herein incorporated by reference.

Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) comprising a promoter operably linked to a polynucleotide of the invention *ex vivo*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, see Belldegrin et al., J. Natl. Cancer Inst., 85:207-216 (1993); Ferrantini et al., Cancer Research, 53:107-1112 (1993); Ferrantini et al., J. Immunology 153: 4604-4615 (1994); Kaido, T., et al., Int. J. Cancer 60: 221-229 (1995); Ogura et al., Cancer Research 50: 5102-5106 (1990); Santodonato, et al., Human Gene Therapy 7:1-10 (1996); Santodonato, et al., Gene Therapy 4:1246-1255

(1997); and Zhang, et al., Cancer Gene Therapy 3: 31-38 (1996)), which are herein incorporated by reference. In one embodiment, the cells which are engineered are arterial cells. The arterial cells may be reintroduced into the patient through direct injection to the artery, the tissues surrounding the artery, or through catheter injection.

5       As discussed in more detail below, the polynucleotide constructs can be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, and the like). The polynucleotide constructs may be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

10       In one embodiment, the polynucleotide of the invention is delivered as a naked polynucleotide. The term "naked" polynucleotide, DNA or RNA refers to sequences that are free from any delivery vehicle that acts to assist, promote or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the  
15       invention can also be delivered in liposome formulations and lipofectin formulations and the like can be prepared by methods well known to those skilled in the art. Such methods are described, for example, in U.S. Patent Nos. 5,593,972, 5,589,466, and 5,580,859, which are herein incorporated by reference.

20       The polynucleotide vector constructs of the invention used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Appropriate vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; pSVK3, pBPV, pMSG and pSVL available from Pharmacia; and pEF1/V5, pcDNA3.1, and pRc/CMV2 available from Invitrogen. Other suitable vectors will be  
25       readily apparent to the skilled artisan.

30       Any strong promoter known to those skilled in the art can be used for driving the expression of polynucleotide sequence of the invention. Suitable promoters include adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters,

such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the b-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter for the polynucleotides of the invention.

Unlike other gene therapy techniques, one major advantage of introducing  
5 naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct of the invention can be delivered to the interstitial  
10 space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular, fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic  
15 fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by  
20 injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

25 For the naked acid sequence injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 mg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue  
30 site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration.



The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked DNA  
5 constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The naked polynucleotides are delivered by any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, and so-called "gene guns". These  
10 delivery methods are known in the art.

The constructs may also be delivered with delivery vehicles such as viral sequences, viral particles, liposome formulations, lipofectin, precipitating agents, etc. Such methods of delivery are known in the art.

In certain embodiments, the polynucleotide constructs of the invention are  
15 complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate  
20 intracellular delivery of plasmid DNA (Felgner et al., Proc. Natl. Acad. Sci. USA , 84:7413-7416 (1987), which is herein incorporated by reference); mRNA (Malone et al., Proc. Natl. Acad. Sci. USA , 86:6077-6081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs et al., J. Biol. Chem., 265:10189-10192 (1990), which is herein incorporated by reference), in functional  
25 form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner et al., Proc. Natl Acad. Sci. USA ,  
30 84:7413-7416 (1987), which is herein incorporated by reference). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication NO: WO 90/11092 (which is herein incorporated by reference) for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation  
5 of DOTMA liposomes is explained in the literature, see, e.g., Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7417, which is herein incorporated by reference. Similar methods can be used to prepare liposomes from other cationic lipid materials.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily  
10 available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these  
15 materials are well known in the art.

For example, commercially dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be  
20 prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at  
25 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar  
30 vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome-nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger et al., Methods of Immunology, 101:512-527 (1983),

which is herein incorporated by reference. For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous

5 population of unilamellar liposomes. The material to be entrapped is added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and

10 DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include  $\text{Ca}^{2+}$ -EDTA chelation (Papahadjopoulos et al., *Biochim. Biophys. Acta*, 394:483 (1975); Wilson et al., *Cell*, 17:77 (1979)); ether injection (Deamer et al.,

15 *Biochim. Biophys. Acta*, 443:629 (1976); Ostro et al., *Biochem. Biophys. Res. Commun.*, 76:836 (1977); Fraley et al., *Proc. Natl. Acad. Sci. USA*, 76:3348 (1979)); detergent dialysis (Enoch et al., *Proc. Natl. Acad. Sci. USA*, 76:145 (1979)); and reverse-phase evaporation (REV) (Fraley et al., *J. Biol. Chem.*, 255:10431 (1980); Szoka et al., *Proc. Natl. Acad. Sci. USA*, 75:145 (1978); Schaefer-Ridder et al.,

20 *Science*, 215:166 (1982)), which are herein incorporated by reference.

Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

U.S. Patent NO: 5,676,954 (which is herein incorporated by reference) reports

25 on the injection of genetic material, complexed with cationic liposomes carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication NO: WO 94/9469 (which are herein incorporated by reference) provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622,

30 5,580,859, 5,703,055, and international publication NO: WO 94/9469 (which are herein incorporated by reference) provide methods for delivering DNA-cationic lipid complexes to mammals.

In certain embodiments, cells are be engineered, *ex vivo* or *in vivo*, using a retroviral particle containing RNA which comprises a sequence encoding polypeptides of the invention. Retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, Rous sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, R-2, R-AM, PA12, T19-14X, VT-19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, Human Gene Therapy , 1:5-14 (1990), which is incorporated herein by reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO<sub>4</sub> precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include polynucleotide encoding polypeptides of the invention. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express polypeptides of the invention.

In certain other embodiments, cells are engineered, *ex vivo* or *in vivo*, with polynucleotides of the invention contained in an adenovirus vector. Adenovirus can be manipulated such that it encodes and expresses polypeptides of the invention, and at the same time is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. Adenovirus expression is achieved without integration of the viral DNA into the host cell chromosome, thereby alleviating concerns about insertional mutagenesis. Furthermore, adenoviruses have been used as live enteric vaccines for many years with an excellent safety profile (Schwartz et al., Am. Rev. Respir. Dis., 109:233-238 (1974)). Finally, adenovirus mediated gene transfer has been demonstrated in a number of instances including transfer of alpha-1-antitrypsin and CFTR to the lungs of cotton rats (Rosenfeld et al., Science , 252:431-434 (1991);

Rosenfeld et al., Cell, 68:143-155 (1992)). Furthermore, extensive studies to attempt to establish adenovirus as a causative agent in human cancer were uniformly negative (Green et al. Proc. Natl. Acad. Sci. USA, 76:6606 (1979)).

Suitable adenoviral vectors useful in the present invention are described, for example, in Kozarsky and Wilson, Curr. Opin. Genet. Devel., 3:499-503 (1993); Rosenfeld et al., Cell, 68:143-155 (1992); Engelhardt et al., Human Genet. Ther., 4:759-769 (1993); Yang et al., Nature Genet., 7:362-369 (1994); Wilson et al., Nature, 365:691-692 (1993); and U.S. Patent NO: 5,652,224, which are herein incorporated by reference. For example, the adenovirus vector Ad2 is useful and can be grown in human 293 cells. These cells contain the E1 region of adenovirus and constitutively express E1a and E1b, which complement the defective adenoviruses by providing the products of the genes deleted from the vector. In addition to Ad2, other varieties of adenovirus (e.g., Ad3, Ad5, and Ad7) are also useful in the present invention.

Preferably, the adenoviruses used in the present invention are replication deficient. Replication deficient adenoviruses require the aid of a helper virus and/or packaging cell line to form infectious particles. The resulting virus is capable of infecting cells and can express a polynucleotide of interest which is operably linked to a promoter, for example, the HARP promoter of the present invention, but cannot replicate in most cells. Replication deficient adenoviruses may be deleted in one or more of all or a portion of the following genes: E1a, E1b, E3, E4, E2a, or L1 through L5.

In certain other embodiments, the cells are engineered, *ex vivo* or *in vivo*, using an adeno-associated virus (AAV). AAVs are naturally occurring defective viruses that require helper viruses to produce infectious particles (Muzyczka, Curr. Topics in Microbiol. Immunol., 158:97 (1992)). It is also one of the few viruses that may integrate its DNA into non-dividing cells. Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate, but space for exogenous DNA is limited to about 4.5 kb. Methods for producing and using such AAVs are known in the art. See, for example, U.S. Patent Nos. 5,139,941, 5,173,414, 5,354,678, 5,436,146, 5,474,935, 5,478,745, and 5,589,377.

For example, an appropriate AAV vector for use in the present invention will include all the sequences necessary for DNA replication, encapsidation, and host-cell integration. The polynucleotide construct containing polynucleotides of the invention is inserted into the AAV vector using standard cloning methods, such as those found  
5 in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989). The recombinant AAV vector is then transfected into packaging cells which are infected with a helper virus, using any standard technique, including lipofection, electroporation, calcium phosphate precipitation, etc. Appropriate helper viruses include adenoviruses, cytomegaloviruses, vaccinia viruses, or herpes viruses.  
10 Once the packaging cells are transfected and infected, they will produce infectious AAV viral particles which contain the polynucleotide construct of the invention. These viral particles are then used to transduce eukaryotic cells, either *ex vivo* or *in vivo*. The transduced cells will contain the polynucleotide construct integrated into its genome, and will express the desired gene product.

15 Another method of gene therapy involves operably associating heterologous control regions and endogenous polynucleotide sequences (e.g. encoding the polypeptide sequence of interest) via homologous recombination (see, e.g., U.S. Patent NO: 5,641,670, issued June 24, 1997; International Publication NO: WO 96/29411, published September 26, 1996; International Publication NO: WO  
20 94/12650, published August 4, 1994; Koller et al., *Proc. Natl. Acad. Sci. USA*, 86:8932-8935 (1989); and Zijlstra et al., *Nature*, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not normally expressed in the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made, using standard techniques known in the  
25 art, which contain the promoter with targeting sequences flanking the promoter. Suitable promoters are described herein. The targeting sequence is sufficiently complementary to an endogenous sequence to permit homologous recombination of the promoter-targeting sequence with the endogenous sequence. The targeting sequence will be sufficiently near the 5' end of the desired endogenous  
30 polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination.

The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter. The amplified promoter and targeting sequences are digested and ligated together.

The promoter-targeting sequence construct is delivered to the cells, either as naked polynucleotide, or in conjunction with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, whole viruses, lipofection, precipitating agents, etc., described in more detail above. The P promoter-targeting sequence can be delivered by any method, included direct needle injection, intravenous injection, topical administration, catheter infusion, particle accelerators, etc. The methods are described in more detail below.

The promoter-targeting sequence construct is taken up by cells. Homologous recombination between the construct and the endogenous sequence takes place, such that an endogenous sequence is placed under the control of the promoter. The promoter then drives the expression of the endogenous sequence.

The polynucleotides encoding polypeptides of the present invention may be administered along with other polynucleotides encoding other angiogenic proteins. Angiogenic proteins include, but are not limited to, acidic and basic fibroblast growth factors, VEGF-1, epidermal growth factor alpha and beta, platelet-derived endothelial cell growth factor, platelet-derived growth factor, tumor necrosis factor alpha, hepatocyte growth factor, insulin like growth factor, colony stimulating factor, macrophage colony stimulating factor, granulocyte/macrophage colony stimulating factor, and nitric oxide synthase.

Preferably, the polynucleotide encoding a polypeptide of the invention contains a secretory signal sequence that facilitates secretion of the protein. Typically, the signal sequence is positioned in the coding region of the polynucleotide to be expressed towards or at the 5' end of the coding region. The signal sequence may be homologous or heterologous to the polynucleotide of interest and may be

homologous or heterologous to the cells to be transfected. Additionally, the signal sequence may be chemically synthesized using methods known in the art.

Any mode of administration of any of the above-described polynucleotides constructs can be used so long as the mode results in the expression of one or more  
5 molecules in an amount sufficient to provide a therapeutic effect. This includes direct needle injection, systemic injection, catheter infusion, biolistic injectors, particle accelerators (i.e., "gene guns"), gelfoam sponge depots, other commercially available depot materials, osmotic pumps (e.g., Alza minipumps), oral or suppository solid (tablet or pill) pharmaceutical formulations, and decanting or topical applications  
10 during surgery. For example, direct injection of naked calcium phosphate-precipitated plasmid into rat liver and rat spleen or a protein-coated plasmid into the portal vein has resulted in gene expression of the foreign gene in the rat livers. (Kaneda et al., Science, 243:375 (1989)).

A preferred method of local administration is by direct injection. Preferably, a  
15 recombinant molecule of the present invention complexed with a delivery vehicle is administered by direct injection into or locally within the area of arteries. Administration of a composition locally within the area of arteries refers to injecting the composition centimeters and preferably, millimeters within arteries.

Another method of local administration is to contact a polynucleotide  
20 construct of the present invention in or around a surgical wound. For example, a patient can undergo surgery and the polynucleotide construct can be coated on the surface of tissue inside the wound or the construct can be injected into areas of tissue inside the wound.

Therapeutic compositions useful in systemic administration, include  
25 recombinant molecules of the present invention complexed to a targeted delivery vehicle of the present invention. Suitable delivery vehicles for use with systemic administration comprise liposomes comprising ligands for targeting the vehicle to a particular site.

Preferred methods of systemic administration, include intravenous injection,  
30 aerosol, oral and percutaneous (topical) delivery. Intravenous injections can be performed using methods standard in the art. Aerosol delivery can also be performed using methods standard in the art (see, for example, Stribling et al., Proc. Natl. Acad.



Sci. USA , 189:11277-11281 (1992), which is incorporated herein by reference). Oral delivery can be performed by complexing a polynucleotide construct of the present invention to a carrier capable of withstanding degradation by digestive enzymes in the gut of an animal. Examples of such carriers, include plastic capsules or tablets, such as those known in the art. Topical delivery can be performed by mixing a polynucleotide construct of the present invention with a lipophilic reagent (e.g., DMSO) that is capable of passing into the skin.

Determining an effective amount of substance to be delivered can depend upon a number of factors including, for example, the chemical structure and biological activity of the substance, the age and weight of the animal, the precise condition requiring treatment and its severity, and the route of administration. The frequency of treatments depends upon a number of factors, such as the amount of polynucleotide constructs administered per dose, as well as the health and history of the subject. The precise amount, number of doses, and timing of doses will be determined by the attending physician or veterinarian. Therapeutic compositions of the present invention can be administered to any animal, preferably to mammals and birds. Preferred mammals include humans, dogs, cats, mice, rats, rabbits sheep, cattle, horses and pigs, with humans being particularly

#### 20 Biological Activities

The polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides or polypeptides, or agonists or antagonists could be used to treat the associated disease.

#### Immune Activity

The polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called

hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins),  
5 or infectious. Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating or detecting deficiencies or disorders of  
10 hematopoietic cells. A polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to:  
15 blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia,  
20 or hemoglobinuria.

Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotides or polypeptides, or agonists or  
25 antagonists of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention that can decrease hemostatic or thrombolytic activity could be used  
30 to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response  
5 leading to the destruction of the host tissue. Therefore, the administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

10 Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid,  
15 Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly  
20 allergic asthma) or other respiratory problems, may also be treated by a polynucleotides or polypeptides, or agonists or antagonists of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotides or polypeptides, or agonists or antagonists of the present  
25 invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polynucleotides or polypeptides, or  
30 agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat

5 inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over

10 production of cytokines (e.g., TNF or IL-1.)

#### **Hyperproliferative Disorders**

A polynucleotides or polypeptides, or agonists or antagonists of the invention can be used to treat or detect hyperproliferative disorders, including neoplasms. A

15 polynucleotides or polypeptides, or agonists or antagonists of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

20 For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an

25 immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotides or polypeptides, or agonists or antagonists of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast,

30 digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and

peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotides or polypeptides, or agonists or antagonists of the present invention.

- 5 Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

10

#### **Cardiovascular Disorders**

- Polynucleotides or polypeptides, or agonists or antagonists of the invention may be used to treat cardiovascular disorders, including peripheral artery disease, such as limb ischemia.
- 15

- Cardiovascular disorders include cardiovascular abnormalities, such as arterio-arterial fistula, arteriovenous fistula, cerebral arteriovenous malformations, congenital heart defects, pulmonary atresia, and Scimitar Syndrome. Congenital heart defects include aortic coarctation, cor triatriatum, coronary vessel anomalies, crisscross heart, dextrocardia, patent ductus arteriosus, Ebstein's anomaly, Eisenmenger complex, hypoplastic left heart syndrome, levocardia, tetralogy of fallot, transposition of great vessels, double outlet right ventricle, tricuspid atresia, persistent truncus arteriosus, and heart septal defects, such as aortopulmonary septal defect, endocardial cushion defects, Lutembacher's Syndrome, trilogly of Fallot, ventricular heart septal defects.
- 20

- Cardiovascular disorders also include heart disease, such as arrhythmias, carcinoid heart disease, high cardiac output, low cardiac output, cardiac tamponade, endocarditis (including bacterial), heart aneurysm, cardiac arrest, congestive heart failure, congestive cardiomyopathy, paroxysmal dyspnea, cardiac edema, heart hypertrophy, congestive cardiomyopathy, left ventricular hypertrophy, right ventricular hypertrophy, post-infarction heart rupture, ventricular septal rupture, heart valve diseases, myocardial diseases, myocardial ischemia, pericardial effusion, pericarditis (including constrictive and tuberculous), pneumopericardium,
- 25
- 30

postpericardiotomy syndrome, pulmonary heart disease, rheumatic heart disease, ventricular dysfunction, hyperemia, cardiovascular pregnancy complications, Scimitar Syndrome, cardiovascular syphilis, and cardiovascular tuberculosis.

Arrhythmias include sinus arrhythmia, atrial fibrillation, atrial flutter,  
5 bradycardia, extrasystole, Adams-Stokes Syndrome, bundle-branch block, sinoatrial block, long QT syndrome, parasystole, Lown-Ganong-Levine Syndrome, Mahaim-type pre-excitation syndrome, Wolff-Parkinson-White syndrome, sick sinus syndrome, tachycardias, and ventricular fibrillation. Tachycardias include paroxysmal tachycardia, supraventricular tachycardia, accelerated idioventricular  
10 rhythm, atrioventricular nodal reentry tachycardia, ectopic atrial tachycardia, ectopic junctional tachycardia, sinoatrial nodal reentry tachycardia, sinus tachycardia, Torsades de Pointes, and ventricular tachycardia.

Heart valve disease include aortic valve insufficiency, aortic valve stenosis, hear murmurs, aortic valve prolapse, mitral valve prolapse, tricuspid valve prolapse,  
15 mitral valve insufficiency, mitral valve stenosis, pulmonary atresia, pulmonary valve insufficiency, pulmonary valve stenosis, tricuspid atresia, tricuspid valve insufficiency, and tricuspid valve stenosis.

Myocardial diseases include alcoholic cardiomyopathy, congestive cardiomyopathy, hypertrophic cardiomyopathy, aortic subvalvular stenosis,  
20 pulmonary subvalvular stenosis, restrictive cardiomyopathy, Chagas cardiomyopathy, endocardial fibroelastosis, endomyocardial fibrosis, Kearns Syndrome, myocardial reperfusion injury, and myocarditis.

Myocardial ischemias include coronary disease, such as angina pectoris, coronary aneurysm, coronary arteriosclerosis, coronary thrombosis, coronary  
25 vasospasm, myocardial infarction and myocardial stunning.

Cardiovascular diseases also include vascular diseases such as aneurysms, angiodyplasia, angiomatosis, bacillary angiomatosis, Hippel-Lindau Disease, Klippel-Trenaunay-Weber Syndrome, Sturge-Weber Syndrome, angioneurotic edema, aortic diseases, Takayasu's Arteritis, aortitis, Leriche's Syndrome, arterial occlusive  
30 diseases, arteritis, enarteritis, polyarteritis nodosa, cerebrovascular disorders, diabetic angiopathies, diabetic retinopathy, embolisms, thrombosis, erythromelalgia, hemorrhoids, hepatic veno-occlusive disease, hypertension, hypotension, ischemia,

peripheral vascular diseases, phlebitis, pulmonary veno-occlusive disease, Raynaud's disease, CREST syndrome, retinal vein occlusion, Scimitar syndrome, superior vena cava syndrome, telangiectasia, atacia telangiectasia, hereditary hemorrhagic telangiectasia, varicocele, varicose veins, varicose ulcer, vasculitis, and venous  
5 insufficiency.

Aneurysms include dissecting aneurysms, false aneurysms, infected aneurysms, ruptured aneurysms, aortic aneurysms, cerebral aneurysms, coronary aneurysms, heart aneurysms, and iliac aneurysms.

Arterial occlusive diseases include arteriosclerosis, intermittent claudication,  
10 carotid stenosis, fibromuscular dysplasias, mesenteric vascular occlusion, Moyamoya disease, renal artery obstruction, retinal artery occlusion, and thromboangiitis obliterans.

Cerebrovascular disorders include carotid artery diseases, cerebral amyloid angiopathy, cerebral aneurysm, cerebral anoxia, cerebral arteriosclerosis, cerebral  
15 arteriovenous malformation, cerebral artery diseases, cerebral embolism and thrombosis, carotid artery thrombosis, sinus thrombosis, Wallenberg's syndrome, cerebral hemorrhage, epidural hematoma, subdural hematoma, subaraxhnoid hemorrhage, cerebral infarction, cerebral ischemia (including transient), subclavian steal syndrome, periventricular leukomalacia, vascular headache, cluster headache,  
20 migraine, and vertebrobasilar insufficiency.

Embolisms include air embolisms, amniotic fluid embolisms, cholesterol embolisms, blue toe syndrome, fat embolisms, pulmonary embolisms, and thromboembolisms. Thrombosis include coronary thrombosis, hepatic vein thrombosis, retinal vein occlusion, carotid artery thrombosis, sinus thrombosis,  
25 Wallenberg's syndrome, and thrombophlebitis.

Ischemia includes cerebral ischemia, ischemic colitis, compartment syndromes, anterior compartment syndrome, myocardial ischemia, reperfusion injuries, and peripheral limb ischemia. Vasculitis includes aortitis, arteritis, Behcet's Syndrome, Churg-Strauss Syndrome, mucocutaneous lymph node syndrome,  
30 thromboangiitis obliterans, hypersensitivity vasculitis, Schoenlein-Henoch purpura, allergic cutaneous vasculitis, and Wegener's granulomatosis.

Polynucleotides or polypeptides, or agonists or antagonists of the invention, are especially effective for the treatment of critical limb ischemia and coronary disease. As shown in the Examples, administration of polynucleotides and polypeptides of the invention to an experimentally induced ischemia rabbit hindlimb  
5 may restore blood pressure ratio, blood flow, angiographic score, and capillary density.

Polypeptides may be administered using any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, biolistic injectors, particle  
10 accelerators, gelfoam sponge depots, other commercially available depot materials, osmotic pumps, oral or suppository solid pharmaceutical formulations, decanting or topical applications during surgery, aerosol delivery. Such methods are known in the art. Polypeptides of the invention may be administered as part of a pharmaceutical composition, described in more detail below. Methods of delivering polynucleotides  
15 of the invention are described in more detail herein.

#### **Anti-Angiogenesis Activity**

The naturally occurring balance between endogenous stimulators and inhibitors of angiogenesis is one in which inhibitory influences predominate.  
20 Rastinejad *et al.*, *Cell* 56:345-355 (1989). In those rare instances in which neovascularization occurs under normal physiological conditions, such as wound healing, organ regeneration, embryonic development, and female reproductive processes, angiogenesis is stringently regulated and spatially and temporally delimited. Under conditions of pathological angiogenesis such as that characterizing  
25 solid tumor growth, these regulatory controls fail. Unregulated angiogenesis becomes pathologic and sustains progression of many neoplastic and non-neoplastic diseases. A number of serious diseases are dominated by abnormal neovascularization including solid tumor growth and metastases, arthritis, some types of eye disorders, and psoriasis. See, e.g., reviews by Moses *et al.*, *Biotech.* 9:630-634 (1991); Folkman  
30 *et al.*, *N. Engl. J. Med.*, 333:1757-1763 (1995); Auerbach *et al.*, *J. Microvasc. Res.* 29:401-411 (1985); Folkman, *Advances in Cancer Research*, eds. Klein and Weinhouse, Academic Press, New York, pp. 175-203 (1985); Patz, *Am. J.*



*Ophthalmol.* 94:715-743 (1982); and Folkman *et al.*, *Science* 221:719-725 (1983). In a number of pathological conditions, the process of angiogenesis contributes to the disease state. For example, significant data have accumulated which suggest that the growth of solid tumors is dependent on angiogenesis. Folkman and Klagsbrun,  
5 *Science* 235:442-447 (1987).

The present invention provides for treatment of diseases or disorders associated with neovascularization by administration of the polynucleotides or polypeptides, or agonists or antagonists of the invention. Malignant and metastatic conditions which can be treated with the polynucleotides and polypeptides, or  
10 agonists or antagonists of the invention include, but are not limited to, malignancies, solid tumors, and cancers described herein and otherwise known in the art (for a review of such disorders, see Fishman *et al.*, *Medicine*, 2d Ed., J. B. Lippincott Co., Philadelphia (1985)).

Ocular disorders associated with neovascularization which can be treated with  
15 the polynucleotides or polypeptides or agonists or antagonists of the invention include, but are not limited to: neovascular glaucoma, diabetic retinopathy, retinoblastoma, retrolental fibroplasia, uveitis, retinopathy of prematurity macular degeneration, corneal graft neovascularization, as well as other eye inflammatory diseases, ocular tumors and diseases associated with choroidal or iris  
20 neovascularization. See, e.g., reviews by Waltman *et al.*, *Am. J. Ophthalmol.* 85:704-710 (1978) and Gartner *et al.*, *Surv. Ophthalmol.* 22:291-312 (1978).

Additionally, disorders which can be treated with the polynucleotides and polypeptides of the present invention (including agonist and/or antagonists) include, but are not limited to, hemangioma, arthritis, psoriasis, angiofibroma, atherosclerotic  
25 plaques, delayed wound healing, granulations, hemophilic joints, hypertrophic scars, nonunion fractures, Osler-Weber syndrome, pyogenic granuloma, scleroderma, trachoma, and vascular adhesions.

Moreover, disorders and/or states, which can be treated with be treated with polynucleotides or polypeptides or agonists or antagonists of the present invention,  
30 but are not limited to, solid tumors, blood born tumors such as leukemias, tumor metastasis, Kaposi's sarcoma, benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas, rheumatoid arthritis,

psoriasis, ocular angiogenic diseases, for example, diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, and uveitis, delayed wound healing, endometriosis, vasculogenesis, granulations, hypertrophic scars (keloids), nonunion  
5 fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischemic limb angiogenesis, Osler-Webber Syndrome, plaque neovascularization, telangiectasia, hemophilic joints, angiofibroma fibromuscular dysplasia, wound granulation, Crohn's disease, atherosclerosis, birth control agent by preventing vascularization  
10 required for embryo implantation controlling menstruation, diseases that have angiogenesis as a pathologic consequence such as cat scratch disease (*Rochelium quintosa*), ulcers (*Helicobacter pylori*), Bartonellosis and bacillary angiomatosis.

15       **Diseases at the Cellular Level**

Diseases associated with increased cell survival or the inhibition of apoptosis that could be treated or detected by the polynucleotides or polypeptides and/or antagonists or agonists of the invention, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors,  
20 including, but not limited to colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, breast cancer, prostate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune disorders  
25 (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation, graft v. host disease, acute graft rejection, and chronic graft rejection. In preferred  
30 embodiments, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention are used to inhibit growth, progression, and/or metastasis of cancers, in particular those listed above.

Additional diseases or conditions associated with increased cell survival that could be treated or detected by the polynucleotides or polypeptides, or agonists or antagonists of the invention, include, but are not limited to, progression, and/or metastases of malignancies and related disorders such as leukemia (including acute leukemias (e.g., acute lymphocytic leukemia, acute myelocytic leukemia (including myeloblastic, promyelocytic, myelomonocytic, monocytic, and erythroleukemia)) and chronic leukemias (e.g., chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia)), polycythemia vera, lymphomas (e.g., Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors including, but not limited to, sarcomas and carcinomas such as fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma, and retinoblastoma.

Diseases associated with increased apoptosis that could be treated or detected by the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, include AIDS; neurodegenerative disorders (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration and brain tumor or prior associated disease); autoimmune disorders (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) myelodysplastic syndromes (such as aplastic anemia), graft v. host disease, ischemic

injury (such as that caused by myocardial infarction, stroke and reperfusion injury), liver injury (e.g., hepatitis related liver injury, ischemia/reperfusion injury, cholestosis (bile duct injury) and liver cancer); toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia.

5

#### **Wound Healing and Epithelial Cell Proliferation**

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, for therapeutic purposes, for example, to stimulate  
10 epithelial cell proliferation and basal keratinocytes for the purpose of wound healing, and to stimulate hair follicle production and healing of dermal wounds. Polynucleotides or polypeptides, as well as agonists or antagonists of the invention, may be clinically useful in stimulating wound healing including surgical wounds, excisional wounds, deep wounds involving damage of the dermis and epidermis, eye  
15 tissue wounds, dental tissue wounds, oral cavity wounds, diabetic ulcers, dermal ulcers, cubitus ulcers, arterial ulcers, venous stasis ulcers, burns resulting from heat exposure or chemicals, and other abnormal wound healing conditions such as uremia, malnutrition, vitamin deficiencies and complications associated with systemic treatment with steroids, radiation therapy and antineoplastic drugs and  
20 antimetabolites. Polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote dermal reestablishment subsequent to dermal loss

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to increase the adherence of skin grafts to a wound bed and  
25 to stimulate re-epithelialization from the wound bed. The following are a non-exhaustive list of grafts that polynucleotides or polypeptides, agonists or antagonists of the invention, could be used to increase adherence to a wound bed: autografts, artificial skin, allografts, autodermic graft, autoepdermic grafts, avacular grafts, Blair-Brown grafts, bone graft, brephoplastic grafts, cutis graft, delayed graft, dermic graft,  
30 epidermic graft, fascia graft, full thickness graft, heterologous graft, xenograft, homologous graft, hyperplastic graft, lamellar graft, mesh graft, mucosal graft, Ollier-Thiersch graft, omenpal graft, patch graft, pedicle graft, penetrating graft, split skin

graft, thick split graft. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, can be used to promote skin strength and to improve the appearance of aged skin.

It is believed that the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, will also produce changes in hepatocyte proliferation, and epithelial cell proliferation in the lung, breast, pancreas, stomach, small intestine, and large intestine. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could promote proliferation of epithelial cells such as sebocytes, hair follicles, hepatocytes, type II pneumocytes, mucin-producing goblet cells, and other epithelial cells and their progenitors contained within the skin, lung, liver, and gastrointestinal tract. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may promote proliferation of endothelial cells, keratinocytes, and basal keratinocytes.

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to reduce the side effects of gut toxicity that result from radiation, chemotherapy treatments or viral infections. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may have a cytoprotective effect on the small intestine mucosa. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may also stimulate healing of mucositis (mouth ulcers) that result from chemotherapy and viral infections.

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could further be used in full regeneration of skin in full and partial thickness skin defects, including burns, (i.e., repopulation of hair follicles, sweat glands, and sebaceous glands), treatment of other skin defects such as psoriasis. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat epidermolysis bullosa, a defect in adherence of the epidermis to the underlying dermis which results in frequent, open and painful blisters by accelerating reepithelialization of these lesions. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to treat gastric and duodenal ulcers and help heal by scar formation of the mucosal lining and regeneration of glandular mucosa and duodenal mucosal lining more rapidly.

Inflammatory bowel diseases, such as Crohn's disease and ulcerative colitis, are diseases which result in destruction of the mucosal surface of the small or large intestine, respectively. Thus, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote the resurfacing of the mucosal surface to aid more rapid healing and to prevent progression of inflammatory bowel disease. Treatment with the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, is expected to have a significant effect on the production of mucus throughout the gastrointestinal tract and could be used to protect the intestinal mucosa from injurious substances that are ingested or following surgery.

10 The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat diseases associate with the under expression of the polynucleotides of the invention.

Moreover, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to prevent and heal damage to the lungs due to various pathological states. A growth factor such as the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, which could stimulate proliferation and differentiation and promote the repair of alveoli and brochiolar epithelium to prevent or treat acute or chronic lung damage. For example, emphysema, which results in the progressive loss of aveoli, and inhalation injuries, i.e., resulting from smoke inhalation and burns, that cause necrosis of the bronchiolar epithelium and alveoli could be effectively treated using the polynucleotides or polypeptides, and/or agonists or antagonists of the invention. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to stimulate the proliferation of and differentiation of type II pneumocytes, which may help treat or prevent disease such as hyaline membrane diseases, such as infant respiratory distress syndrome and bronchopulmonary displasia, in premature infants.

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could stimulate the proliferation and differentiation of hepatocytes and, thus, could be used to alleviate or treat liver diseases and pathologies such as fulminant liver failure caused by cirrhosis, liver damage caused by viral hepatitis and toxic substances (i.e., acetaminophen, carbon tetraholoride and other hepatotoxins known in the art).

In addition, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used treat or prevent the onset of diabetes mellitus. In patients with newly diagnosed Types I and II diabetes, where some islet cell function remains, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to maintain the islet function so as to alleviate, delay or prevent permanent manifestation of the disease. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used as an auxiliary in islet cell transplantation to improve or promote islet cell function.

#### 10 Infectious Disease

A polypeptide or polynucleotide and/or agonist or antagonist of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, polypeptide or polynucleotide and/or agonist or antagonist of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections

(e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide, and/or agonist or antagonist of the present invention  
5 can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g.,  
10 Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales,  
15 Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis,  
20 uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis,  
25 Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide and/or agonist or antagonist of the present invention can be used to treat or detect any of these symptoms or diseases.

30 Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, the following families: Amebiasis,



Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide and/or agonist or antagonist of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide and/or agonist or antagonist of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

### **Regeneration**

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds  
5 includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide and/or agonist or antagonist of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this  
10 method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g.,  
15 Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide and/or agonist or antagonist of the present invention.

### **Chemotaxis**

20 A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then  
25 fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted  
30 to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured

location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may inhibit chemotactic activity. These molecules  
5 could also be used to treat disorders. Thus, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention could be used as an inhibitor of chemotaxis.

### **Binding Activity**

10 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or  
15 small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural  
20 receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell  
25 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

30 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving

competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product  
5 mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a  
10 sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

Additionally, the receptor to which a polypeptide of the invention binds can be identified by numerous methods known to those of skill in the art, for example, ligand  
15 panning and FACS sorting (Coligan, et al., Current Protocols in Immun., 1(2), Chapter 5, (1991)). For example, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the polypeptides, for example, NIH3T3 cells which are known to contain multiple receptors for the FGF family proteins, and SC-3 cells, and a cDNA library created from this RNA is divided  
20 into pools and used to transfect COS cells or other cells that are not responsive to the polypeptides. Transfected cells which are grown on glass slides are exposed to the polypeptide of the present invention, after they have been labelled. The polypeptides can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase.

25 Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and retransfected using an iterative sub-pooling and re-screening process, eventually yielding a single clones that encodes the putative receptor.

As an alternative approach for receptor identification, the labeled polypeptides  
30 can be photoaffinity linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE analysis and exposed to X-ray film. The labeled complex containing the receptors of the

polypeptides can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the genes encoding the putative receptors.

- 5           Moreover, the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling") may be employed to modulate the activities of polypeptides of the invention thereby effectively generating agonists and antagonists of polypeptides of the invention. See generally, U.S. Patent Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, and  
10 5,837,458, and Patten, P. A., et al., *Curr. Opinion Biotechnol.* 8:724-33 (1997); Harayama, S. *Trends Biotechnol.* 16(2):76-82 (1998); Hansson, L. O., et al., *J. Mol. Biol.* 287:265-76 (1999); and Lorenzo, M. M. and Blasco, R. *Biotechniques* 24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference). In one embodiment, alteration of polynucleotides and corresponding  
15 polypeptides of the invention may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments into a desired polynucleotide sequence of the invention molecule by homologous, or site-specific, recombination. In another embodiment, polynucleotides and corresponding polypeptides of the invention may be altered by being subjected to random mutagenesis by error-prone  
20 PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of the polypeptides of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules. In preferred embodiments, the heterologous molecules are  
25 family members. In further preferred embodiments, the heterologous molecule is a growth factor such as, for example, platelet-derived growth factor (PDGF), insulin-like growth factor (IGF-I), transforming growth factor (TGF)-alpha, epidermal growth factor (EGF), fibroblast growth factor (FGF), TGF-beta, bone morphogenetic protein (BMP)-2, BMP-4, BMP-5, BMP-6, BMP-7, activins A and B,  
30 decapentaplegic(dpp), 60A, OP-2, dorsalin, growth differentiation factors (GDFs), nodal, MIS, inhibin-alpha, TGF-beta1, TGF-beta2, TGF-beta3, TGF-beta5, and glial-derived neurotrophic factor (GDNF).

Other preferred fragments are biologically active fragments of the polypeptides of the invention. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide. The biological activity of the fragments may include an improved desired activity, or a  
5 decreased undesirable activity.

Additionally, this invention provides a method of screening compounds to identify those which modulate the action of the polypeptide of the present invention. An example of such an assay comprises combining a mammalian fibroblast cell, a the polypeptide of the present invention, the compound to be screened and 3[H]  
10 thymidine under cell culture conditions where the fibroblast cell would normally proliferate. A control assay may be performed in the absence of the compound to be screened and compared to the amount of fibroblast proliferation in the presence of the compound to determine if the compound stimulates proliferation by determining the uptake of 3[H] thymidine in each case. The amount of fibroblast cell proliferation is  
15 measured by liquid scintillation chromatography which measures the incorporation of 3[H] thymidine. Both agonist and antagonist compounds may be identified by this procedure.

In another method, a mammalian cell or membrane preparation expressing a receptor for a polypeptide of the present invention is incubated with a labeled  
20 polypeptide of the present invention in the presence of the compound. The ability of the compound to enhance or block this interaction could then be measured. Alternatively, the response of a known second messenger system following interaction of a compound to be screened and the receptor is measured and the ability of the compound to bind to the receptor and elicit a second messenger response is  
25 measured to determine if the compound is a potential agonist or antagonist. Such second messenger systems include but are not limited to, cAMP guanylate cyclase, ion channels or phosphoinositide hydrolysis.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring  
30 about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptides of the invention from

suitably manipulated cells or tissues. Therefore, the invention includes a method of identifying compounds which bind to the polypeptides of the invention comprising the steps of: (a) incubating a candidate binding compound with the polypeptide; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with the polypeptide, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Also, one could identify molecules bind a polypeptide of the invention experimentally by using the beta-pleated sheet regions contained in the polypeptide sequence of the protein. Accordingly, specific embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, the amino acid sequence of each beta pleated sheet regions in a disclosed polypeptide sequence. Additional embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, any combination or all of contained in the polypeptide sequences of the invention. Additional preferred embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, the amino acid sequence of each of the beta pleated sheet regions in one of the polypeptide sequences of the invention. Additional embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, any combination or all of the beta pleated sheet regions in one of the polypeptide sequences of the invention.

### **Drug Screening**

Further contemplated is the use of the polypeptides of the present invention, or the polynucleotides encoding these polypeptides, to screen for molecules which modify the activities of the polypeptides of the present invention. Such a method would include contacting the polypeptide of the present invention with a selected compound(s) suspected of having antagonist or agonist activity, and assaying the activity of these polypeptides following binding.

This invention is particularly useful for screening therapeutic compounds by using the polypeptides of the present invention, or binding fragments thereof, in any of a variety of drug screening techniques. The polypeptide or fragment employed in

such a test may be affixed to a solid support, expressed on a cell surface, free in solution, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed  
5 cells in competitive binding assays. One may measure, for example, the formulation of complexes between the agent being tested and a polypeptide of the present invention.

Thus, the present invention provides methods of screening for drugs or any other agents which affect activities mediated by the polypeptides of the present  
10 invention. These methods comprise contacting such an agent with a polypeptide of the present invention or a fragment thereof and assaying for the presence of a complex between the agent and the polypeptide or a fragment thereof, by methods well known in the art. In such a competitive binding assay, the agents to screen are typically labeled. Following incubation, free agent is separated from that present in bound  
15 form, and the amount of free or uncomplexed label is a measure of the ability of a particular agent to bind to the polypeptides of the present invention.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the polypeptides of the present invention, and is described in great detail in European Patent Application 84/03564,  
20 published on September 13, 1984, which is incorporated herein by reference herein. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with polypeptides of the present invention and washed. Bound polypeptides are then detected by methods well known in the art.  
25 Purified polypeptides are coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies may be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding polypeptides of the present  
30 invention specifically compete with a test compound for binding to the polypeptides or fragments thereof. In this manner, the antibodies are used to detect the presence of



any peptide which shares one or more antigenic epitopes with a polypeptide of the invention.

#### **Antisense And Ribozyme (Antagonists)**

5           In specific embodiments, antagonists according to the present invention are nucleic acids corresponding to the sequences contained in SEQ ID NO:X, or the complementary strand thereof, and/or to nucleotide sequences contained a deposited clone. In one embodiment, antisense sequence is generated internally by the organism, in another embodiment, the antisense sequence is separately administered  
10 (see, for example, O'Connor, Neurochem., 56:560 (1991). Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Antisense technology can be used to control gene expression through antisense DNA or RNA, or through triple-helix formation. Antisense techniques are discussed for example, in Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense  
15 Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance, Lee et al., Nucleic Acids Research, 6:3073 (1979); Cooney et al., Science, 241:456 (1988); and Dervan et al., Science, 251:1300 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA.

20           For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of the receptor. The antisense  
25 RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into receptor polypeptide.

          In one embodiment, the antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector or a portion thereof, is transcribed, producing an antisense nucleic acid (RNA) of the  
30 invention. Such a vector would contain a sequence encoding the antisense nucleic acid of the invention. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA.

Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding a polypeptide of the invention, or fragments thereof, can be by any promoter known in the art to act  
5 in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, *Nature*, 29:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell*, 22:787-797 (1980), the herpes thymidine promoter (Wagner et al., *Proc. Natl. Acad. Sci. U.S.A.*,  
10 78:1441-1445 (1981), the regulatory sequences of the metallothionein gene (Brinster et al., *Nature*, 296:39-42 (1982)), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a gene of interest. However, absolute complementarity, although preferred, is not required. A sequence  
15 "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double stranded antisense nucleic acids of the invention, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of  
20 complementarity and the length of the antisense nucleic acid. Generally, the larger the hybridizing nucleic acid, the more base mismatches with a RNA sequence of the invention it may contain and still form a stable duplex (or triplex as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

25 Oligonucleotides that are complementary to the 5' end of the message, *e.g.*, the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., *Nature*,  
30 372:333-335 (1994). Thus, oligonucleotides complementary to either the 5' - or 3' - non-translated, non-coding regions of a polynucleotide sequence of the invention could be used in an antisense approach to inhibit translation of endogenous mRNA.

Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the  
5 5' -, 3' - or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

The polynucleotides of the invention can be DNA or RNA or chimeric  
10 mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell  
15 membrane (see, e.g., Letsinger et al., Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci., 84:648-652 (1987); PCT Publication NO: WO88/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication NO: WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., BioTechniques, 6:958-976 (1988))  
20 or intercalating agents. (See, e.g., Zon, Pharm. Res., 5:539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety  
25 which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine,  
30 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine,

5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v),  
5 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

10 In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

15 In yet another embodiment, the antisense oligonucleotide is an a-anomeric oligonucleotide. An a-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run parallel to each other (Gautier et al., Nucl. Acids Res., 15:6625-6641 (1987)). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., Nucl. Acids Res.,  
20 15:6131-6148 (1987)), or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215:327-330 (1987)).

Polynucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.): As examples,  
25 phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (Nucl. Acids Res., 16:3209 (1988)), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., Proc. Natl. Acad. Sci. U.S.A., 85:7448-7451 (1988)), etc.

While antisense nucleotides complementary to the coding region sequence of  
30 the invention could be used, those complementary to the transcribed untranslated region are most preferred.

Potential antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al, Science, 247:1222-1225 (1990). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy mRNAs  
5 corresponding to the polynucleotides of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well  
10 known in the art and is described more fully in Haseloff and Gerlach, Nature, 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within each nucleotide sequence disclosed in the sequence listing. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA corresponding to the polynucleotides of the invention; i.e., to  
15 increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

As in the antisense approach, the ribozymes of the invention can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the polynucleotides of the invention *in vivo*.  
20 DNA constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive promoter, such as, for example, pol III or pol II promoter, so that transfected cells will produce sufficient quantities of  
25 the ribozyme to destroy endogenous messages and inhibit translation. Since ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Antagonist/agonist compounds may be employed to inhibit the cell growth and proliferation effects of the polypeptides of the present invention on neoplastic  
30 cells and tissues, i.e. stimulation of angiogenesis of tumors, and, therefore, retard or prevent abnormal cellular growth and proliferation, for example, in tumor formation or growth.

The antagonist/agonist may also be employed to prevent hyper-vascular diseases, and prevent the proliferation of epithelial lens cells after extracapsular cataract surgery. Prevention of the mitogenic activity of the polypeptides of the present invention may also be desirous in cases such as restenosis after balloon  
5 angioplasty.

The antagonist/agonist may also be employed to prevent the growth of scar tissue during wound healing.

The antagonist/agonist may also be employed to treat the diseases described herein.

10

#### Other Activities

The polypeptide of the present invention, as a result of the ability to stimulate vascular endothelial cell growth, may be employed in treatment for stimulating re-vascularization of ischemic tissues due to various disease conditions such as  
15 thrombosis, arteriosclerosis, and other cardiovascular conditions. These polypeptide may also be employed to stimulate angiogenesis and limb regeneration, as discussed above.

The polypeptide may also be employed for treating wounds due to injuries, burns, post-operative tissue repair, and ulcers since they are mitogenic to various cells  
20 of different origins, such as fibroblast cells and skeletal muscle cells, and therefore, facilitate the repair or replacement of damaged or diseased tissue.

The polypeptide of the present invention may also be employed stimulate neuronal growth and to treat and prevent neuronal damage which occurs in certain neuronal disorders or neuro-degenerative conditions such as Alzheimer's disease,  
25 Parkinson's disease, and AIDS-related complex. The polypeptide of the invention may have the ability to stimulate chondrocyte growth, therefore, they may be employed to enhance bone and periodontal regeneration and aid in tissue transplants or bone grafts.

The polypeptide of the present invention may be also be employed to prevent  
30 skin aging due to sunburn by stimulating keratinocyte growth.

The polypeptide of the invention may also be employed for preventing hair loss, since FGF family members activate hair-forming cells and promotes melanocyte

growth. Along the same lines, the polypeptides of the present invention may be employed to stimulate growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines.

The polypeptide of the invention may also be employed to maintain organs  
5 before transplantation or for supporting cell culture of primary tissues.

The polypeptide of the present invention may also be employed for inducing tissue of mesodermal origin to differentiate in early embryos.

The polypeptide or polynucleotides and/or agonist or antagonists of the  
10 present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue,  
15 pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, polypeptides or polynucleotides and/or agonist or antagonists of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

Polypeptide or polynucleotides and/or agonist or antagonists of the present  
20 invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

25 Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

### 30 **Other Preferred Embodiments**

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

5 Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

10 Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

15 Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

20 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

25 Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.



A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under  
5 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which  
10 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide  
15 sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of  
20 at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule  
30 comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X  
5 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and  
10 determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence  
15 selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

20 A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X  
25 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological  
30 sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of

positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence  
5 at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

10 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino  
15 acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted  
20 protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the  
25 amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in  
30 the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA

clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of

the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1  
 5 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an  
 10 isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

15

### Examples

#### Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector.  
 20 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being  
 25 isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

<u>Vector Used to Construct Library</u>		<u>Corresponding Deposited</u>
<u>Plasmid</u>		
30	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA



pSport1	pSport1
pCMVSPORT 2.0	pCMVSPORT 2.0
pCMVSPORT 3.0	pCMVSPORT 3.0
pCR <sup>®</sup> 2.1	pCR <sup>®</sup> 2.1

- 5        Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are
- 10       commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3
- 15       primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.
- 20       Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University,
- 25       NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR<sup>®</sup>2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al.,
- 30       Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ - $\gamma$ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25  $\mu\text{l}$  of reaction

mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A<sup>+</sup> RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific

to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5    **Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

10    **Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P<sup>32</sup> using the rediprime™ DNA labeling  
15    system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

20    Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed  
25    according to standard procedures.

**Example 4: Chromosomal Mapping of the Polynucleotides**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This  
30    primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and

hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

#### **Example 5: Bacterial Expression of a Polypeptide**

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp<sup>r</sup>), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan<sup>r</sup>). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.<sup>600</sup>) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1

mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The

origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

**Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without  
5 mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive  
10 Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4  
15 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using  
20 a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant  $A_{280}$  monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the  
25 above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL  
assays.



**Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector  
5 contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak  
10 *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such  
15 as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

20 Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP)  
25 to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a  
30 commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

5       The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by  
10   gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method  
15   described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the  
20   transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

25       After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell  
30   culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then

resuspended in a microcentrifuge tube containing 200  $\mu$ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

5 To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and  
10 cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5  $\mu$ Ci of  $^{35}$ S-methionine and 5  $\mu$ Ci  $^{35}$ S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

15 Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

#### **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian  
20 cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient  
25 transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include,  
30 for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109),

pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

5           Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

          The transfected gene can also be amplified to express large amounts of the  
10    encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker  
15    is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the  
20    production of proteins.

          Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the  
25    CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

30           Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be  
5 modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

10 The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

15 Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five  $\mu\text{g}$  of the expression plasmid pC6 a pC4 is cotransfected with 0.5  $\mu\text{g}$  of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are  
20 seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in  
25 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1  $\mu\text{M}$ , 2  $\mu\text{M}$ , 5  $\mu\text{M}$ , 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200  $\mu\text{M}$ . Expression of the desired gene product is analyzed,  
30 for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

#### **Example 9: Protein Fusions**

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see  
5 also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create  
10 chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

15 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion  
20 can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will  
25 not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

30

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC  
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAA  
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTACATGCGTGGT  
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG  
5 ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA  
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT  
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA  
ACCCCATCGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCGAGAAC  
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG  
10 GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT  
GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAATAACAAGACCACGCCT  
CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG  
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA  
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG  
15 GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

**Example 10: Production of an Antibody from a Polypeptide**

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) As one example of such methods, cells  
20 expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

25 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell  
30 Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in

any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56 degrees C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

5           The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as  
10 described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

          Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method  
15 makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody  
20 whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

          It will be appreciated that Fab and F(ab')<sub>2</sub> and other fragments of the  
25 antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic  
30 chemistry.

          For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced



using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

**Example 11: Production Of Secreted Protein For High-Throughput Screening**

**Assays**

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45

minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37 degrees C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl<sub>2</sub> (anhyd); 0.00130 mg/L CuSO<sub>4</sub>·5H<sub>2</sub>O; 0.050 mg/L of Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O; 0.417 mg/L of FeSO<sub>4</sub>·7H<sub>2</sub>O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl<sub>2</sub>; 48.84 mg/L of MgSO<sub>4</sub>; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO<sub>3</sub>; 62.50 mg/L of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O; 71.02 mg/L of Na<sub>2</sub>HPO<sub>4</sub>; 4320 mg/L of ZnSO<sub>4</sub>·7H<sub>2</sub>O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H<sub>2</sub>O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H<sub>2</sub>O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H<sub>2</sub>O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H<sub>2</sub>O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319

- mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B<sub>12</sub>; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of
- 5 Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for
- 10 endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37 degrees C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

- 15 On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

- It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the
- 20 polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

25 **Example 12: Construction of GAS Reporter Construct**

- One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The
- 30 binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u> <u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATs</u>	<u>GAS(elements) or ISRE</u>
	<u>IFN family</u>						
5	IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
	IL-10	+	?	?	-	1,3	
	<u>gp130 family</u>						
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	IL-11(Pleiotrophic)	?	+	?	?	1,3	
	OnM(Pleiotrophic)	?	+	+	?	1,3	
	LIF(Pleiotrophic)	?	+	+	?	1,3	
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3	
	IL-12(Pleiotrophic)	+	-	+	+	1,3	
	<u>g-C family</u>						
20	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
25	IL-15	?	+	?	+	5	GAS
	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
30	GM-CSF (myeloid)	-	-	+	-	5	GAS
	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
35	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	
40	CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

20 5':CTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCT  
25 AGGCTTTTGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter  
30 molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not  
5 contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance  
10 gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS  
15 with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-  
20 2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

**Example 13: High-Throughput Screening Assay for T-cell Activity.**

25 The following protocol is used to assess T-cell activity by identifying factors, determining whether supernate containing a polypeptide of the invention proliferates and/or differentiates T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used  
30 in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately  
5 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells  
10 containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

15 During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37 degrees C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

20 The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing polypeptides of the invention and/or induced polypeptides of the invention as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and  
25 resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells  
30 into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).



After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20 degrees C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4 degrees C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as, stable transfected cells, which would be apparent to those of skill in the art.

#### **Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by determining whether polypeptides of the invention proliferates and/or differentiates myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 1 mM MgCl<sub>2</sub>, and 675 uM CaCl<sub>2</sub>. Incubate at 37 degrees C for 45 min.

- 5 Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37 degrees C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

- 10 These cells are tested by harvesting  $1 \times 10^8$  cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200 ul cells per well in the 96-well plate (or  $1 \times 10^5$  cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example

- 15 11. Incubate at 37 degrees C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

20 **Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.**

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the

25 EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

- 30 The EGR1 gene expression is activated during this treatment. Thus, by stably

transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

- 5           Add 200  $\mu$ l of the cell suspension to each well of 96-well plate (equivalent to  $1 \times 10^5$  cells/well). Add 50  $\mu$ l supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ $\mu$ l of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the  
10           supernatant according to Example 17.

**Example 16: High-Throughput Screening Assay for T-cell Activity**

- NF- $\kappa$ B (Nuclear Factor  $\kappa$ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and  
15           CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- $\kappa$ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- $\kappa$ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

- 20           In non-stimulated conditions, NF- $\kappa$ B is retained in the cytoplasm with I- $\kappa$ B (Inhibitor  $\kappa$ B). However, upon stimulation, I- $\kappa$ B is phosphorylated and degraded, causing NF- $\kappa$ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- $\kappa$ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

- 25           Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- $\kappa$ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- $\kappa$ B would be useful in treating diseases. For example, inhibitors of NF- $\kappa$ B could be used to treat those  
30           diseases related to the acute or chronic activation of NF- $\kappa$ B, such as rheumatoid arthritis.

To construct a vector containing the NF- $\kappa$ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- $\kappa$ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5' : GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC  
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

10 5' : GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following

15 sequence:

5' : CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGGACTTTCC  
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC  
ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA  
20 CTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTA  
TTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA  
GCTT:3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2- promoter plasmid (Clontech) with this NF- $\kappa$ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- $\kappa$ B/SV40/SEAP cassette is removed from the above NF- $\kappa$ B/SEAP vector using restriction enzymes

30 SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- $\kappa$ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- $\kappa$ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

#### **Example 17: Assay for SEAP Activity**

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15  $\mu$ l of 2.5x dilution buffer into Optiplates containing 35  $\mu$ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65 degrees C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50  $\mu$ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50  $\mu$ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

#### **Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5

19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

**Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used

instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37 degrees C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to  $2-5 \times 10^6$  cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37 degrees C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to  $1 \times 10^6$  cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca<sup>++</sup> concentration.

#### **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase**

##### **Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine



Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also  
5 membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and  
10 non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of  
15 activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from  
20 Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 µl of cell culture grade type I collagen (50 µg/ml), gelatin (2%) or polylysine (50 µg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or  
25 calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen  
30 Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in

5 Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 2 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4 degrees C. The plate is then placed in a

10 vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4

15 degrees C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by

20 determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

25 The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sub>2</sub><sup>+</sup> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the

30 components gently and preincubate the reaction mix at 30 degrees C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37 degrees C for 20  
5 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37 degrees C for one hour. Wash the well as above.  
10 Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

15

**Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation  
20 (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other  
25 phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G  
30 plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any

of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4 degrees C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants  
5 obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in  
10 place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive  
15 incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

**Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide**  
20

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest  
25 in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95 degrees C for 30 seconds; 60-120 seconds at 52-58 degrees C; and 60-120 seconds at 70 degrees C, using buffer solutions described in Sidransky et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre  
30 Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring

suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton et al., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenin deoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

25

**Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample**

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

30

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in

- 5 Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled  
10 water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

- 15 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).  
20 Interpolate the concentration of the polypeptide in the sample using the standard curve.

### **Example 23: Formulating a Polypeptide**

- The secreted polypeptide composition will be formulated and dosed in a  
25 fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

- 30 As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 ug/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will

be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Compositions of the invention are also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include suitable polymeric materials (such as, for example, semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules), suitable hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, and sparingly soluble derivatives (such as, for example, a sparingly soluble salt).

Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al., *Biopolymers* 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (Langer et al., *J. Biomed. Mater. Res.* 15:167-277 (1981), and Langer, *Chem. Tech.* 12:98-105 (1982)), ethylene vinyl acetate (Langer et al., *Id.*) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988).

Sustained-release compositions also include liposomally entrapped compositions of the invention (*see generally*, Langer, *Science* 249:1527-1533 (1990); Treat et al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-

Berestein and Fidler (eds.), Liss, New York, pp. 317 -327 and 353-365 (1989)).

Liposomes containing XXX polypeptide may be prepared by methods known per se:

DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. (USA) 82:3688-3692 (1985);

Hwang et al., Proc. Natl. Acad. Sci. (USA) 77:4030-4034 (1980); EP 52,322; EP

5 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal XXX polypeptide therapy.

10 In yet an additional embodiment, the compositions of the invention are delivered by way of a pump (*see* Langer, *supra*; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)).

Other controlled release systems are discussed in the review by Langer  
15 (*Science* 249:1527-1533 (1990)).

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and  
20 concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both.  
25 Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

30 The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as



phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as

5 polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

10 The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile.

15 Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for  
20 example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

25 The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the  
30 agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

The compositions of the invention may be administered alone or in combination with other therapeutic agents. Therapeutic agents that may be administered in combination with the compositions of the invention, include but not limited to, other members of the TNF family, chemotherapeutic agents, antibiotics, 5 steroidal and non-steroidal anti-inflammatories, conventional immunotherapeutic agents, cytokines and/or growth factors. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations in which the combined agents are administered together as a therapeutic mixture, and also procedures in which the 10 combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given first, followed by the second.

In one embodiment, the compositions of the invention are administered in 15 combination with members of the TNF family. TNF, TNF-related or TNF-like molecules that may be administered with the compositions of the invention include, but are not limited to, soluble forms of TNF-alpha, lymphotoxin-alpha (LT-alpha, also known as TNF-beta), LT-beta (found in complex heterotrimer LT-alpha2-beta), OPGL, FasL, CD27L, CD30L, CD40L, 4-1BBL, DcR3, OX40L, TNF-gamma 20 (International Publication No. WO 96/14328), AIM-I (International Publication No. WO 97/33899), endokine-alpha (International Publication No. WO 98/07880), TR6 (International Publication No. WO 98/30694), OPG, and neutrokin-alpha (International Publication No. WO 98/18921, OX40, and nerve growth factor (NGF), and soluble forms of Fas, CD30, CD27, CD40 and 4-1BB, TR2 (International 25 Publication No. WO 96/34095), DR3 (International Publication No. WO 97/33904), DR4 (International Publication No. WO 98/32856), TR5 (International Publication No. WO 98/30693), TR6 (International Publication No. WO 98/30694), TR7 (International Publication No. WO 98/41629), TRANK, TR9 (International Publication No. WO 98/56892), TR10 (International Publication No. WO 98/54202), 30 312C2 (International Publication No. WO 98/06842), and TR12, and soluble forms CD154, CD70, and CD153.

Conventional nonspecific immunosuppressive agents, that may be administered in combination with the compositions of the invention include, but are not limited to, steroids, cyclosporine, cyclosporine analogs, cyclophosphamide methylprednisone, prednisone, azathioprine, FK-506, 15-deoxyspergualin, and other  
5 immunosuppressive agents that act by suppressing the function of responding T cells.

In a further embodiment, the compositions of the invention are administered in combination with an antibiotic agent. Antibiotic agents that may be administered with the compositions of the invention include, but are not limited to, tetracycline, metronidazole, amoxicillin, beta-lactamases, aminoglycosides, macrolides,  
10 quinolones, fluoroquinolones, cephalosporins, erythromycin, ciprofloxacin, and streptomycin.

In an additional embodiment, the compositions of the invention are administered alone or in combination with an anti-inflammatory agent. Anti-inflammatory agents that may be administered with the compositions of the invention  
15 include, but are not limited to, glucocorticoids and the nonsteroidal anti-inflammatories, aminoarylcarboxylic acid derivatives, arylacetic acid derivatives, arylbutyric acid derivatives, arylcarboxylic acids, arylpropionic acid derivatives, pyrazoles, pyrazolones, salicylic acid derivatives, thiazinecarboxamides, e-acetamidocaproic acid, S-adenosylmethionine, 3-amino-4-hydroxybutyric acid,  
20 amixetrine, bendazac, benzydamine, bucolome, difenpiramide, ditazol, emorfazone, guaiazulene, nabumetone, nimesulide, orgotein, oxaceprol, paranyline, perisoxal, pifoxime, proquazone, proxazole, and tenidap.

In another embodiment, compositions of the invention are administered in combination with a chemotherapeutic agent. Chemotherapeutic agents that may be  
25 administered with the compositions of the invention include, but are not limited to, antibiotic derivatives (e.g., doxorubicin, bleomycin, daunorubicin, and dactinomycin); antiestrogens (e.g., tamoxifen); antimetabolites (e.g., fluorouracil, 5-FU, methotrexate, floxuridine, interferon alpha-2b, glutamic acid, plicamycin, mercaptopurine, and 6-thioguanine); cytotoxic agents (e.g., carmustine, BCNU,  
30 lomustine, CCNU, cytosine arabinoside, cyclophosphamide, estramustine, hydroxyurea, procarbazine, mitomycin, busulfan, cis-platin, and vincristine sulfate); hormones (e.g., medroxyprogesterone, estramustine phosphate sodium, ethinyl

estradiol, estradiol, megestrol acetate, methyltestosterone, diethylstilbestrol diphosphate, chlorotrianisene, and testolactone); nitrogen mustard derivatives (e.g., mephalen, chorambucil, mechlorethamine (nitrogen mustard) and thiotepa); steroids and combinations (e.g., bethamethasone sodium phosphate); and others (e.g.,  
5 dicarbazine, asparaginase, mitotane, vincristine sulfate, vinblastine sulfate, and etoposide).

In an additional embodiment, the compositions of the invention are administered in combination with cytokines. Cytokines that may be administered with the compositions of the invention include, but are not limited to, IL2, IL3, IL4,  
10 IL5, IL6, IL7, IL10, IL12, IL13, IL15, anti-CD40, CD40L, IFN-gamma and TNF-alpha.

In an additional embodiment, the compositions of the invention are administered in combination with angiogenic proteins. Angiogenic proteins that may be administered with the compositions of the invention include, but are not limited to,  
15 Glioma Derived Growth Factor (GDGF), as disclosed in European Patent Number EP-399816; Platelet Derived Growth Factor-A (PDGF-A), as disclosed in European Patent Number EP-682110; Platelet Derived Growth Factor-B (PDGF-B), as disclosed in European Patent Number EP-282317; Placental Growth Factor (PIGF), as disclosed in International Publication Number WO 92/06194; Placental Growth  
20 Factor-2 (PIGF-2), as disclosed in Hauser et al., Growth Factors, 4:259-268 (1993); Vascular Endothelial Growth Factor (VEGF), as disclosed in International Publication Number WO 90/13649; Vascular Endothelial Growth Factor-A (VEGF-A), as disclosed in European Patent Number EP-506477; Vascular Endothelial Growth Factor-2 (VEGF-2), as disclosed in International Publication Number WO 96/39515;  
25 Vascular Endothelial Growth Factor B-186 (VEGF-B186), as disclosed in International Publication Number WO 96/26736; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/02543; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/07832; and Vascular Endothelial Growth Factor-E  
30 (VEGF-E), as disclosed in German Patent Number DE19639601. The above mentioned references are incorporated herein by reference herein.

In an additional embodiment, the compositions of the invention are administered in combination with Fibroblast Growth Factors. Fibroblast Growth Factors that may be administered with the compositions of the invention include, but are not limited to, FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, 5 FGF-9, FGF-10, FGF-11, FGF-12, FGF-13, FGF-14, and FGF-15.

In additional embodiments, the compositions of the invention are administered in combination with other therapeutic or prophylactic regimens, such as, for example, radiation therapy.

10 Example 24: Method of Treating Decreased Levels of the Polypeptide

The present invention relates to a method for treating an individual in need of an increased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an agonist of the invention (including polypeptides of the 15 invention). Moreover, it will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering 20 to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on 25 administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

The present invention also relates to a method of treating an individual in need of a decreased level of a polypeptide of the invention in the body comprising 30 administering to such an individual a composition comprising a therapeutically effective amount of an antagonist of the invention (including polypeptides and antibodies of the invention).

In one example, antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

10 **Example 26: Method of Treatment Using Gene Therapy-Ex Vivo**

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1 using primers and having appropriate restriction sites and initiation/stop codons, if necessary. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine

sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto  
5 agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the  
10 gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media,  
15 containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no  
20 selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or  
25 after having been grown to confluence on cytodex 3 microcarrier beads.

#### **Example 27: Gene Therapy Using Endogenous Genes Corresponding To Polynucleotides of the Invention**

Another method of gene therapy according to the present invention involves  
30 operably associating the endogenous polynucleotide sequence of the invention with a promoter via homologous recombination as described, for example, in U.S. Patent NO: 5,641,670, issued June 24, 1997; International Publication NO: WO 96/29411,

published September 26, 1996; International Publication NO: WO 94/12650,  
published August 4, 1994; Koller et al., *Proc. Natl. Acad. Sci. USA*, 86:8932-8935  
(1989); and Zijlstra et al., *Nature*, 342:435-438 (1989). This method involves the  
activation of a gene which is present in the target cells, but which is not expressed in  
5 the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made which contain a promoter and targeting  
sequences, which are homologous to the 5' non-coding sequence of endogenous  
polynucleotide sequence, flanking the promoter. The targeting sequence will be  
sufficiently near the 5' end of the polynucleotide sequence so the promoter will be  
10 operably linked to the endogenous sequence upon homologous recombination. The  
promoter and the targeting sequences can be amplified using PCR. Preferably, the  
amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends.  
Preferably, the 3' end of the first targeting sequence contains the same restriction  
enzyme site as the 5' end of the amplified promoter and the 5' end of the second  
15 targeting sequence contains the same restriction site as the 3' end of the amplified  
promoter.

The amplified promoter and the amplified targeting sequences are digested  
with the appropriate restriction enzymes and subsequently treated with calf intestinal  
phosphatase. The digested promoter and digested targeting sequences are added  
20 together in the presence of T4 DNA ligase. The resulting mixture is maintained under  
conditions appropriate for ligation of the two fragments. The construct is size  
fractionated on an agarose gel then purified by phenol extraction and ethanol  
precipitation.

In this Example, the polynucleotide constructs are administered as naked  
25 polynucleotides via electroporation. However, the polynucleotide constructs may also  
be administered with transfection-facilitating agents, such as liposomes, viral  
sequences, viral particles, precipitating agents, etc. Such methods of delivery are  
known in the art.

Once the cells are transfected, homologous recombination will take place  
30 which results in the promoter being operably linked to the endogenous polynucleotide  
sequence. This results in the expression of polynucleotide corresponding to the



polynucleotide in the cell. Expression may be detected by immunological staining, or any other method known in the art.

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in DMEM + 10% fetal calf serum. Exponentially growing or early stationary  
5 phase fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of electroporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl, 5 mM KCl, 0.7 mM Na<sub>2</sub> HPO<sub>4</sub>, 6 mM dextrose). The cells are recentrifuged, the  
10 supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension contains approximately  $3 \times 10^6$  cells/ml. Electroporation should be performed immediately following resuspension.

Plasmid DNA is prepared according to standard techniques. For example, to  
15 construct a plasmid for targeting to the locus corresponding to the polynucleotide of the invention, plasmid pUC18 (MBI Fermentas, Amherst, NY) is digested with HindIII. The CMV promoter is amplified by PCR with an XbaI site on the 5' end and a BamHI site on the 3' end. Two non-coding sequences are amplified via PCR: one non-coding sequence (fragment 1) is amplified with a HindIII site at the 5' end and an  
20 Xba site at the 3' end; the other non-coding sequence (fragment 2) is amplified with a BamHI site at the 5' end and a HindIII site at the 3' end. The CMV promoter and the fragments (1 and 2) are digested with the appropriate enzymes (CMV promoter - XbaI and BamHI; fragment 1 - XbaI; fragment 2 - BamHI) and ligated together. The resulting ligation product is digested with HindIII, and ligated with the HindIII-  
25 digested pUC18 plasmid.

Plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad). The final DNA concentration is generally at least 120 µg/ml. 0.5 ml of the cell suspension (containing approximately  $1.5 \times 10^6$  cells) is then added to the cuvette, and the cell suspension and DNA solutions are gently mixed. Electroporation is  
30 performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at 960 µF and 250-300 V, respectively. As voltage increases, cell survival decreases, but the percentage of surviving cells that stably incorporate the introduced DNA into their

genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer  
5 pipette. The cells are added directly to 10 ml of prewarmed nutrient media (DMEM with 15% calf serum) in a 10 cm dish and incubated at 37 degree C. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a further 16-24 hours.

The engineered fibroblasts are then injected into the host, either alone or after  
10 having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product. The fibroblasts can then be introduced into a patient as described above.

**Example 28: Method of Treatment Using Gene Therapy - In Vivo**

15 Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter  
20 or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata et al., Cardiovasc. Res. 35(3):470-479 (1997); Chao et al., Pharmacol. Res. 35(6):517-522 (1997); Wolff, Neuromuscul. Disord. 7(5):314-318  
25 (1997); Schwartz et al., Gene Ther. 3(5):405-411 (1996); Tsurumi et al., Circulation 94(12):3281-3290 (1996) (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The  
30 polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are

free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in  
5 Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain  
10 sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to  
15 provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and  
20 connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of  
25 the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely  
30 differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after

preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

5

**Example 29: Transgenic Animals.**

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, *e.g.*,  
10 baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (*i.e.*,  
15 polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus  
20 mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, *e.g.*, Ulmer et al., Science  
25 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

30 Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to

quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their  
5 cells, *i.e.*, mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory  
10 sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the  
15 endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science  
20 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening  
25 may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse  
30 transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate  
5 lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to  
10 produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of  
15 polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

**Example 30: Knock-Out Animals.**

20 Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional  
25 polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in  
30 the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in

research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (*e.g.*, see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (*i.e.*, animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (*e.g.*, lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, *e.g.*, by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, *e.g.*, in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, *e.g.*, genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).



When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing  
5 for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders  
10 associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

**Example 31. Isolation of antibody fragments directed against polypeptides of the invention from a library of scFvs.**

15 Naturally occurring V-genes isolated from human PBLs are constructed into a large library of antibody fragments which contain reactivities against a polypeptide having the amino acid sequence of SEQ ID NO:Y to which the donor may or may not have been exposed (see e.g., U.S. Patent 5,885,793 incorporated herein in its entirety by reference).

**Rescue of the library.**

A library of scFvs is constructed from the RNA of human PBLs as described in WO92/01047. To rescue phage displaying antibody fragments, approximately  $10^9$  *E. coli* harboring the phagemid are used to inoculate 50 ml of 2x TY containing 1% glucose and 100 micrograms/ml of ampicillin (2xTY-AMP-GLU) and grown to an O.D. of 0.8 with shaking. Five ml of this culture is used to inoculate 50 ml of 2xTY-AMP-GLU,  $2 \times 10^8$  TU of delta gene 3 helper (M13 delta gene III, see WO92/01047) are added and the culture incubated at 37°C for 45 minutes without shaking and then at 37°C for 45 minutes with shaking. The culture is centrifuged at 4000 r.p.m. for 10 min. and the pellet resuspended in 2 liters of 2x TY containing 100 micrograms/ml ampicillin and 50 micrograms/ml kanamycin and grown overnight. Phage are prepared as described in WO92/01047.

M13 delta gene III is prepared as follows: M13 delta gene III helper phage does not encode gene III protein, hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13 delta gene III particles are made by growing the helper phage in cells harboring a pUC19 derivative supplying the wild type gene III protein during phage morphogenesis. The culture is incubated for 1 hour at 37°C without shaking and then for a further hour at 37°C with shaking. Cells were spun down (IEC-Centra 8, 4000 revs/min for 10 min), resuspended in 300 ml 2x TY broth containing 100 micrograms ampicillin/ml and 25 micrograms kanamycin/ml (2x TY-AMP-KAN) and grown overnight, shaking at 37°C. Phage particles are purified and concentrated from the culture medium by two PEG-precipitations (Sambrook et al., 1990), resuspended in 2 ml PBS and passed through a 0.45 micrometer filter (Minisart NML; Sartorius) to give a final concentration of approximately  $10^{13}$  transducing units/ml (ampicillin-resistant clones).

**Panning the Library.**

Immunotubes (Nunc) are coated overnight in PBS with 4 ml of either 100 micrograms/ml or 10 micrograms/ml of a polypeptide of the present invention. Tubes are blocked with 2% Marvel-PBS for 2 hours at 37°C and then washed 3 times in PBS. Approximately  $10^{13}$  TU of phage is applied to the tube and

incubated for 30 minutes at room temperature tumbling on an over and under turntable and then left to stand for another 1.5 hours. Tubes are washed 10 times with PBS 0.1% Tween-20 and 10 times with PBS. Phage are eluted by adding 1 ml of 100 mM triethylamine and rotating 15 minutes on an under and over  
5 turntable after which the solution is immediately neutralized with 0.5 ml of 1.0M Tris-HCl, pH 7.4. Phage are then used to infect 10 ml of mid-log *E. coli* TG1 by incubating eluted phage with bacteria for 30 minutes at 37°C. The *E. coli* are then plated on TYE plates containing 1% glucose and 100 micrograms/ml ampicillin. The resulting bacterial library is then rescued with delta gene 3 helper phage as  
10 described above to prepare phage for a subsequent round of selection. This process is then repeated for a total of 4 rounds of affinity purification with tube-washing increased to 20 times with PBS, 0.1% Tween-20 and 20 times with PBS for rounds 3 and 4.

#### Characterization of Binders.

15 Eluted phage from the third and fourth rounds of selection are used to infect *E. coli* HB 2151 and soluble scFv is produced (Marks, et al., 1991) from single colonies for assay. ELISAs are performed with microtiter plates coated with either 10 picograms/ml of the polypeptide of the present invention in 50 mM bicarbonate pH 9.6. Clones positive in ELISA are further characterized by PCR fingerprinting  
20 (see e.g., WO92/01047) and then by sequencing.

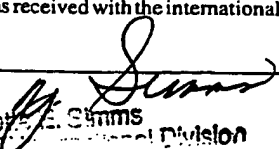
It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above  
25 teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence  
30 listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>129</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit  <u>203105</u>	Accession Number  <u>August 13, 1998</u>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable)	
This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer  Yvonne E. Simms International Division	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

**ATCC Deposit No. 203105****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**ATCC Deposit No. 203105****DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

289

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>129</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <div style="text-align: center;"><u>203081</u></div>	Accession Number <div style="text-align: center;"><u>July 30, 1998</u></div>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
<b>EUROPE</b> In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application         </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">           Authorized officer  <div style="text-align: center;">              Yvette L. Simms              PCT International Division           </div> </div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on:         </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">           Authorized officer         </div>
---	--

**ATCC Deposit No. 203081****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.



**ATCC Deposit No. 203081****DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

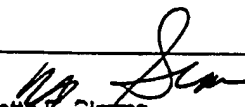
**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>131</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit  <div style="text-align: center;"><u>209090</u></div>	Accession Number  <div style="text-align: center;"><u>June 5, 1997</u></div>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
EUROPE In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application       </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">         Authorized officer  <div style="text-align: center;">             Yvette E. Strams            PCT International Division         </div> </div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on:       </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">         Authorized officer       </div>
--	--

**ATCC Deposit No. 209090****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**ATCC Deposit No. 209090****DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

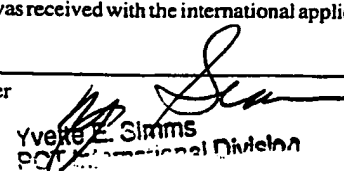
**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>131</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>209215</u>	Accession Number <u>August 21, 1997</u>
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
<b>EUROPE</b> In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<b>For receiving Office use only</b> <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer  Yvette E. Slims PCT International Division	<b>For International Bureau use only</b> <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
--	--

**ATCC Deposit No. 209215****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**ATCC Deposit No. 209215****DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>133</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> <span style="float: right;">Further deposits are identified on an additional sheet <input type="checkbox"/></span>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution <i>(including postal code and country)</i> <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit  <u>203027</u>	Accession Number  <u>June 26, 1998</u>
<b>C. ADDITIONAL INDICATIONS</b> <i>(leave blank if not applicable)</i> <span style="float: right;">This information is continued on an additional sheet <input type="checkbox"/></span>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> <i>(if the indications are not for all designated States)</i>	
<b>EUROPE</b> In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later <i>(specify the general nature of the indications e.g., "Accession Number of Deposit")</i>	
<div style="border: 1px solid black; padding: 5px;"> <b>For receiving Office use only</b>  <input checked="" type="checkbox"/> This sheet was received with the international application    Authorized officer <u>Yvette E. Simms</u>  <u>PCT International Division</u> </div>	<div style="border: 1px solid black; padding: 5px;"> <b>For International Bureau use only</b>  <input type="checkbox"/> This sheet was received by the International Bureau on:    Authorized officer _____ </div>



**ATCC Deposit No. 203027****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**ATCC Deposit No. 203027****DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

***What Is Claimed Is:***

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
  - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
  - (f) a polynucleotide which is a variant of SEQ ID NO:X;
  - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
  - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
  - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- 5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 20 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 25 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
9. A recombinant host cell produced by the method of claim 8.
- 30 10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - 5 (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
  - (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included  
10 in ATCC Deposit No:Z;
  - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - 15 (g) a variant of SEQ ID NO:Y;
  - (h) an allelic variant of SEQ ID NO:Y; or
  - (i) a species homologue of the SEQ ID NO:Y.
12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-  
20 terminus or the N-terminus.
13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
- 25 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
15. A method of making an isolated polypeptide comprising:
- (a) culturing the recombinant host cell of claim 14 under conditions such that  
30 said polypeptide is expressed; and
  - (b) recovering said polypeptide.

16. The polypeptide produced by claim 15.
17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount  
5 of the polypeptide of claim 11 or the polynucleotide of claim 1.
18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or absence of a mutation in the polynucleotide of  
10 claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
19. A method of diagnosing a pathological condition or a susceptibility to  
15 a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological  
20 condition based on the presence or amount of expression of the polypeptide.
20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
- (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of the  
25 polypeptide.
21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.
22. A method of identifying an activity in a biological assay, wherein the  
30 method comprises:
- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;

- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

1

&lt;110&gt; Human Genome Sciences, Inc. et al.

&lt;120&gt; 52 Human Secreted Proteins

&lt;130&gt; PZ032PCT

&lt;140&gt; Unassigned

&lt;141&gt; 1999-08-24

&lt;150&gt; 60/097,917

&lt;151&gt; 1998-08-25

&lt;150&gt; 60/098,634

&lt;151&gt; 1998-08-31

&lt;160&gt; 170

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 733

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

gggatccgga gccc aaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg	60
aatttcgaggg tgcaccgtca gtcttcctct tcccccaaaa acccaaggac accctcatga	120
tctcccggtac tcctgaggtc acatgcgtgg tgggtggacgt aagccacgaa gacctgagg	180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg	240
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact	300
ggctgaatgg caaggagtac aagtgcagg tctccaacaa agccctccca acccccatcg	360
agaaaacccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc	420
catcccggtg tgagctgacc aagaaccagg tcagcctgac ctgcctgggtc aaaggcttct	480
atccaagcga catcgccgtg gagggtggaga gcaatgggca gccggagaac aactacaaga	540
ccacgcctcc cgtgctggac tccgacggct ccttcttctc ctacagcaag ctcaccgtgg	600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc	660
acaaccacta cagcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc	720
gactctagag gat	733

&lt;210&gt; 2

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; Site

&lt;222&gt; (3)

&lt;223&gt; Xaa equals any of the twenty naturally occurring L-amino acids

&lt;400&gt; 2

Trp	Ser	Xaa	Trp	Ser
1				5

&lt;210&gt; 3

&lt;211&gt; 86

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3



2

gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60  
 cccgaaatat ctgccatctc aattag 86

<210> 4  
 <211> 27  
 <212> DNA  
 <213> Homo sapiens

<400> 4  
 gcggcaagct ttttgcaaag cctaggg 27

<210> 5  
 <211> 271  
 <212> DNA  
 <213> Homo sapiens

<400> 5  
 ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60  
 aaatatctgc catctcaatt agtcagcaac catagtcccc cccctaactc cgcccatccc 120  
 gcccttaact cgcgccagtt cgcgccatc tccgccccat ggctgactaa ttttttttat 180  
 ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240  
 ttttgagggc ctaggctttt gcaaaaagct t 271

<210> 6  
 <211> 32  
 <212> DNA  
 <213> Homo sapiens

<400> 6  
 gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7  
 <211> 31  
 <212> DNA  
 <213> Homo sapiens

<400> 7  
 gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8  
 <211> 12  
 <212> DNA  
 <213> Homo sapiens

<400> 8  
 ggggactttc cc 12

<210> 9  
 <211> 73  
 <212> DNA  
 <213> Homo sapiens

<400> 9  
 gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60  
 ccatctcaat tag 73

<210> 10  
 <211> 256  
 <212> DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

ctcgagggga	ctttcccgga	gactttccgg	ggactttccg	ggactttcca	tctgccatct	60
caattagtca	gcaaccatag	tcccgcacct	aactccgccc	atcccgcgcc	taactccgcc	120
cagttccgcc	cattctccgc	cccatggctg	actaattttt	tttatttatg	cagaggccga	180
ggccgcctcg	gcctctgagc	tattccagaa	gtagttagga	ggcttttttg	gaggcctagg	240
cttttgcaaa	aagctt					256

&lt;210&gt; 11

&lt;211&gt; 3289

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 11

cccgggtcga	cccacgcgtc	cggcgaggac	cgcgtccggc	gcagtcttca	atgagcagcg	60
cggaaactgc	accccagacc	cgagcctgct	gcgcgcccc	tcccagagct	cacctggtgc	120
caggtaacag	gcctggcctc	gccctgtgga	tgatgatggc	cttgcccccg	tgagctacaa	180
cctggccttc	agcaccgcc	cacctccaac	cagcaggatg	cggtgtgga	aggcgtggt	240
ggtgactttg	gccttcatga	gtgtggacat	ctgcgtgacc	acggccatct	atgtcttcag	300
ccacctggac	cgcagcctcc	tggaggacat	ccgccacttc	aacatctttg	actcgtgtgt	360
ggatctcttg	gcagcctgcc	tgtaccgcag	ctgcctgctg	ctggggagcc	accattggtg	420
tggccaagaa	cagtgcgtg	gggccccggc	ggctgcgggc	ctcgtggctg	gtcatcacc	480
tcgtgtgcct	cttcgtgggc	atctatgcc	tgggtgaagct	gctgctcttc	tcagaggtgc	540
gcaggcccat	ccgggacccc	tgggtttggg	ccctgttcgt	gtggacgtac	atttactctg	600
gcgcacccct	cctgctctgg	tggctgctgt	ccaccgtgcg	gccaggcacc	caggccctgg	660
agccaggggc	ggccaccgag	gctgaaggct	tccctgggag	cggccggcca	ccgcccgaac	720
aagcgtctgg	ggccacgctg	cagaagctgc	tctcctacac	caagcccgac	gtggccttcc	780
tcgtggcccg	ctcctcttcc	ctcatcgtgg	cagctctggg	agagaccttc	ctgccctact	840
acacgggccc	cgccattgat	ggcatcgtca	tccagaaaag	catggatcag	ttcagcacgg	900
ctgtcgtcat	cgtgtgcctg	ctggccattg	gcagctcatt	tgccgcagg	attcggggcg	960
gcattttttac	cctcatattt	gccagactga	acattcgctt	tcgaaactgt	ctcttccgct	1020
cactgggtgc	ccaggagaca	agcttctttg	atgagaaccg	cacaggggac	ctcatctccc	1080
gcctgacctc	ggacaccacc	atggtcagcg	acctggtctc	cagaacatca	atgtcttctt	1140
gcggaacaca	gtcaagggtca	cgggcgtggt	ggtcttcatg	ttcagcctct	catggcagct	1200
ctccttggtc	accttcatgg	gcttccccat	catcatgatg	gtgtccaaca	tctacggcaa	1260
gtactacaag	aggctctcca	aagagggtcca	gaatgccttg	gccagagcga	gcaacacggc	1320
ggaggagacc	atcagtgcc	tgaagactgt	ccggagcttc	gccaatgagg	aggaggaggc	1380
agagggtgtac	ctgcggaagc	tgacgaggt	gtacaagctg	aacaggaagg	aggcagctgc	1440
ctacatgtac	tacgtctggg	gcagcgggct	cacactgctg	gtggtccagg	tcagcatcct	1500
ctactacggg	ggccaccttg	tcctctcagg	ccagatgacc	agcggcaacc	tcctgcctt	1560
catcatctac	gagtttgtcc	tgggagattg	tatggagaat	gtctcttcca	gcctgtcccc	1620
cggcaagggtg	acggcccttg	tggggccctc	gggcagtggg	aagagctcct	gtgtcaacat	1680
cctggagaac	ttctaccccc	tggagggggg	ccgggtgctg	ctggacggca	agcccatcag	1740
cgcctacgac	cacaagtact	tgcaccgtgt	gatctccctg	gtgagccagg	agcccggtgt	1800
gttcgcccgc	tccatcacgg	ataacatctc	ctacggcctg	ccactgtgtc	ctttcgagat	1860
ggtggtggag	gcccacacaga	aggccaatgc	ccacggcttc	atcatggaac	tccaggacgg	1920
ctacagcaca	gagacagggg	agaagggcgc	ccagctgtca	ggtggccaga	agcagcgggt	1980
ggcatggccc	gggtctggtg	gcggaacccc	cagtcctcca	tcctggatga	agccaccagc	2040
gctttggatg	ccgagagcga	gtatctgata	cagcaggcca	tccatggcaa	cctgcagaag	2100
cacacggtac	tcctcatcgc	gcaccggctg	agcaccgtgg	agcacgcgca	cctcattgtg	2160
gtgctggaca	agggccgctg	agtgcagcag	ggcaccacc	agcagctgct	ggccaggggc	2220
ggcctctacg	ccaagctggt	gcagcggcag	atgctggggc	ttcagccgcg	cgcagacttc	2280
acagctggcc	acaacgagcc	tgtagccaac	ggcagtcaca	aggcctgatg	ggggggccct	2340
gcttctcccc	gtggggcaga	ggacccgggtg	cctgcctggc	agatgtgccc	acggaggccc	2400
ccagctgccc	tccgagccca	ggcctgcagc	actgaaagac	gacctgccat	gtcccatgga	2460
tcaccgcttc	ctgcactctg	cccctggctc	ctgccccatt	cccagggcac	tccttaccct	2520
tgctgccctg	agccaacgcc	ttcacggacc	tccttagcct	cctaagcaaa	ggtagagctg	2580
cctttttaaa	cctagggtctt	accagggttt	ttactggttg	gtttgaggca	ccccagctca	2640

ctcctagatt	tcaaaaaacct	tttttctaatt	gggagtaatg	gcgggcactt	tcaccaagat	2700
gttcttagaaa	cttctgagcc	aggagtgaat	ggcccttcct	tagtagcctg	ggggatgtcc	2760
agagactagg	cctctccctt	ttaccctctc	agagaagggg	cttcctgtc	ccggaggagg	2820
acacggggaa	cgggattttc	cgtctctccc	tcttgccagc	tctgtgagtc	tggccagggc	2880
gggtagggag	cgtggagggc	atctgtctgc	catcgcccg	tgccaatcta	agccagtc	2940
actgtgaacc	acacgaaacc	tcaactgggg	gagtgagggg	ctggccaggt	ctggaggggc	3000
ctcaggggtg	ccccagccc	ggcaccagc	gctttcgccc	ctcgccacc	caccctggc	3060
tggcagcctc	cctccccaca	cccgccctg	tgctctgctg	tctggaggcc	acgtggatgt	3120
tcatgagatg	cattctcttc	tgtctttggt	ggatgggatg	gtggcaaagc	ccaggatctg	3180
gctttgccag	aggttgcaac	atgttgagag	aaccgggtca	ataaagtgt	ctacctctta	3240
cccataaaaa	aaaaaaaaaa	aaaaaaaaaa	aaagggcggc	cgctctaga		3289

&lt;210&gt; 12

&lt;211&gt; 2342

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 12

ggcagagct	cagatctctc	ctggtacccc	ttccccacgc	ccttagataa	tccatctcaa	60
ttcctcatgc	taattgagga	gctatggctg	caaggcacct	tccaggattt	cacacctaca	120
caaatctcct	ttttctcctt	ttgccttctc	tgcttatggg	atattctgag	tccccacccc	180
caatcactga	cagctggggc	cccttcacac	ccctcacaca	ccacgtatta	agtcagtcac	240
aatckccctt	ctcmtctaac	tgctggattt	gtctttctac	acacacccaa	tgattcacgg	300
ctcttcgggc	tgamctactt	acatggacac	agtccaacgt	gagcttacac	atttcttact	360
tggctatacc	attcttggct	gattcattcc	tgaaamcgg	tcataacctg	gaaactcagc	420
taaacacctt	tccttcaaac	tcagctctct	cagcatgggt	tcaggtaggg	ctgttgccct	480
ccttcacctt	atagcttctg	gactaacatc	catacaaac	aacacagcat	catctaaacc	540
accaatattg	gggtaccatt	tctactcaaa	cttccttcac	atccccaccc	cccttatgtc	600
tcagccgaac	ctaccctaac	ccagcccacg	ccacaatgg	gggacagggt	ccccagtcce	660
tatgtgggtc	tatttttacc	cttgcaactc	ctgtagacca	tcaattctac	accctaatta	720
caaaatcata	tccacctctg	cctggcagaa	gggtttatgc	ttttctggct	cgcctaccat	780
ccacacatcc	ctacacctca	ccaccggatc	ctcttttctt	tccttccatc	caattcctgg	840
cttccccgct	gccaactctg	ctctctatgt	ctccagttta	aagggtcccc	ctggaaaaaa	900
tgtacaacatt	ccctcacctg	tgactggtac	ctgacagcca	ccacaccggg	gcagcaatgg	960
ctaacgggtg	acaaagacaa	tttctttctc	tctccaaaac	caaacagcct	tcatcaactc	1020
cctagccaag	actccctatc	aggcccttac	agggtgcgct	ctggctggca	gttaccctaat	1080
ttgggaaaaa	gaaaataccc	tatcatggta	cctaccttca	cctacaactt	tgctgtcacc	1140
cccagttctc	ttttgtgtga	tacaactgat	atttkgccta	ccagccaact	ggtcaggaac	1200
ttgcaccctg	gtctttcagg	ctccaacatc	caacatccta	ccccctaacc	aaactattct	1260
aattttctgta	gaagcctcta	tctcctcttc	accctaaga	aataaatggg	ctctacatct	1320
catcaccttg	ctaacaggat	taggcacac	tgctgcactt	ggcactggaa	tagcaggcat	1380
aaccacctca	atcacctcat	accaaact	attcacacc	ctttctaaca	ccgtagaaga	1440
tatgcacact	tccattacca	gtctccaacg	acaattagac	ttcctcgtgg	gagtcacct	1500
tcaaaactgg	agagtcctgg	acctcctaac	cactgagaaa	gggggtacct	gcatatacct	1560
ccaggaagaa	tgctgtttct	gtgttaatga	atctggcatt	gttcatatcg	cagttcgtag	1620
gcttcatgac	agggctgcag	agctttgaca	tcaagtgcgt	gactcctggg	ggcaaggatc	1680
atcccttcta	agatggatac	cctgggttgc	cccttccta	ggaccctga	tcttctctt	1740
cctgttacta	atgattgggc	catgcatatt	taaccttgta	tcccgcttca	tttcccaaag	1800
gctgaattgt	tttatccagg	caagcatgca	aaaacacatt	gataatatat	ttcacctttg	1860
ccacgtctaa	taccagagcc	tacgaggaaa	ccattcggaa	gctccagaa	ccaggcccta	1920
atcacaacgc	ccctatccag	cagggaagcag	ccagatgac	aacgacgccc	tttttctctt	1980
ttatactaaa	gtaagaaata	agaatgttag	cccaaactgc	actattttgc	agacccttac	2040
cattttacaa	actggtcaga	gtggaaaatt	ccaccagggc	ctgagctgtg	agaaacatcc	2100
tgtcaggcag	gtcccaggcc	taacccttgg	ctgactaaaa	ttccttcatt	atcagcagcc	2160
aaacacaccg	ccccacccc	attttcacaa	caatcccaga	cctctcctgc	ccgggactgt	2220
aactggtcca	gcctgtaagc	gggaaggggg	ctctggcact	agctgggtacc	ccctctccgc	2280
aggtctttct	cccaataaat	ctgtgttgcc	attgaaaaaa	aaaaaaaaaa	aaaaaaactc	2340
ga						2342

<210> 13  
 <211> 1666  
 <212> DNA  
 <213> Homo sapiens

<400> 13  
 ggtggagttc gcacctccag ctccgggccga tgtggaagct ttggagagct gaagagggcg 60  
 cggcggcgct cggcggcgcg ctcttcctgc tgctcttcgc gctaggggtc cgccagctgc 120  
 tgaagcagag gcggccgatg ggcttcccc cggggccgcc ggggctgcc tttatcggca 180  
 acatctattc cctggcagcc tcatccgagc ttccccatgt ctacatgaga aagcagagcc 240  
 aggtgtacgg agagatcttc agtttagatc ttggaggcat atcaactgtg gttctaaatg 300  
 gctatgatgt agtaaaggaa tgccttggtc atcaaagcga aatttttgca gacagaccat 360  
 gccttccttt attcatgaag atgacaaaaa tgggaggctt actcaattcc agatatggcc 420  
 gaggatgggt tgatcacaga cgattagctg taacacagttt tcgatatttt ggatatggcc 480  
 aaaagtcttt tgaatctaaa atcttggaag aaaccaaatt tttcaatgat gctattgaaa 540  
 catacaaagg tagacctttt gactttaaac agttaataac gaatgctgtt tcaaacataa 600  
 ccaatctgat catttttgga gaacgattca cttatgaaga caccgatttt cagcacatga 660  
 ttgagttatt tagtgaaaat gtggaactag ctgccagtc ctcagctctc ttgtataatg 720  
 cctttccatg gattggcatc ctgccttttg gaaaacatca acagctgttt agaaatgcag 780  
 ctgtagtcta tgattttctc tccagactca ttgaaaaagc ttcagtcaac agaaagcctc 840  
 agctacctca gcattttggt gatgcttatt tagatgagat ggatcaagggt aaaaatgacc 900  
 catcatctac tttctccaaa gaaaacctaa tttctcagt gggtgaaactc atcattgctg 960  
 gaactgaaac tacaaccaat gtgctacggt gggcgattct tttcatggcc ctttatccta 1020  
 atattcaagg acaagttcag aaagagattg atttaattat gggccctaata gggaagcctt 1080  
 cttgggacga caaatgcaaa atgccttata ctgaggcagt tttgcatgaa gttttaagat 1140  
 tctgtaatat agttccatta gggattttcc atgcaacctc tgaagatgca gttgtacgtg 1200  
 gttattccat tcctaaaggc acaacagtaa ttacaaatct ttattctgta cactttgatg 1260  
 aaaagtactg gagagaccca gaagtgttcc atcctgagcg atttctggac agcagtggat 1320  
 attttgccaa gaaggaagct ttggttcctt tttccctagg aagaagacat tgtcttgag 1380  
 aacacttgcc tcggatggaa atgttcttgt tttttacagc attgcttcag aggtttcatt 1440  
 tgcattttcc acatgaacta gttccagatc tgaagcccag gttaggcatg acattgcagc 1500  
 cccaacccta cctcatctgt gctgaaagac gctgaaactg cctgggatgt tttcggaac 1560  
 aagaatgtat atttgctta tccctgaact tggtttaatc aaatcaatgt gtgtattaga 1620  
 ataaaagtca cagcatcaaa aagmcaaaaa aaaaaaaaaa aaaaaa 1666

<210> 14  
 <211> 2027  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (294)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1976)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1981)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1985)  
 <223> n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2021)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 14

```

ggcacgagtt gggagcagct ctgctgctg ggcctcagag aatgaggccg gcgttcgccc      60
tgtgcctcct ctggcaggcg ctctggcccc ggccggggcg cggcgaacac cccactgccc      120
accgtgctgg ctgctcggcc tcggggggcct gctacagcct gcaccacgct accatgaagc      180
ggcaggcggc cgaggaggcc tgcacccctgc gaggtggggc gctcagcacc gtgctgctgg      240
gcgccgagct gcgcgctgtg ctgcgctcc tgcgggcagg cccaggggccc gganggggct      300
ccaaagacct gctgttctgg gtcgcactgg agcgcaggcg ttcccactgc amcctggaga      360
acgagccttt gcgggggttt tcctggctgt cctccgaccc cggcgggtctc gaaagcgaca      420
cgctgcagtg ggtggaggag ccccaacgct cctgcaccgc gcggagatgg gtacttccag      480
gccaccgggtg gggctcagcc cgagctgga aggagatgcg atgccacctg ygcgccaacg      540
ctacctgtgc aagtaccagt ttgaggtctt gtgtcctgctg ccgcgccccg gggccgcctc      600
taacttgagc tatcgcgcg ccttccagct gcacagcgcc gctctggact tcagtccacc      660
tgggaccgag gtgagtgcgc tctgccgggg acagctcccc atctcagtta cttgcatcgc      720
ggacgaaatc ggcgctcgyt gggacaaact ytcggggcgt gtgttggtgc cctgccccgg      780
gaggtacctc ctgctggca aatgcgcaga gctccctaac tgcctagacg acttggggagg      840
ctttgcctgc gaatgtgcta cgggcttcga gctgggggaa gacggccgct cttgtgtgac      900
cagtggggaa ggacagccga cccttggggg gaccgggggtg cccaccaggc gcccgccggc      960
cactgcaacc agccccgtgc cgagagaaac atggccaatc agggctcgacg agaagctggg      1020
agagacacca cttgtccctg aacaagacaa ttcagtaaca tctattcctg agattcctcg      1080
atggggatca cagagcacga tgtctaccct tcaaatgtcc cttcaagccg agtcaaaggc      1140
cactatcacc ccatacagga gcgtgatttc caagtttaat tctacgactt cctctgccac      1200
tcctcaggct ttcgactcct cctctgccgt ggtcttcata tttgtgagca cagcagtagt      1260
agtgttggtg atcttgacca tgacagtact ggggcttgct aagctctgct ttcacgaaag      1320
cccctcttcc cagccaagga aggagtctat gggcccgcgg ggctggagag tgatcctgaa      1380
gcccgctgct ttgggctcca gttctgcaca ttgcacaaac aatgggggtga aagtcgggga      1440
ctgtgatctg cgggacagag cagagggtgc cttgctggcg gagtcccctc ttggctctag      1500
tgatgcatag ggaacaggga gacatgggca ctctgtgtaa cagtttttca cttttgatga      1560
aacggggaac caagaggaa ttacttgtgt aactgacaat ttctgcagaa atcccccttc      1620
ctctaaattc cctttactcc actgaggagc taaatcagaa ctgcacactc cttccctgat      1680
gatagaggaa gtggaagtgc ctttaggatg gtgatactgg gggaccgggt agtgctgggg      1740
agagatattt tcttatgttt attcggagaa tttggagaag tgattgaact tttcaagaca      1800
ttggaaacaa atagaacaca atataattta cattaaaaaa taatttctac caaaatggaa      1860
aggaaatggt ctatgttgtt caggctagga gtatattggt tcgaaatccc agggaaaaaa      1920
ataaaaaataa aaaattaaa gattgttgat aaaaaaaaaa aaaaaggggc gccgcncatg      1980
ngggnccaag ctttacgtac gcgggcatgc gacgtcaagc ncttcca                      2027

```

&lt;210&gt; 15

&lt;211&gt; 2334

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2278)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2290)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 15

```

gggagtgtgg ctgcagaacc cagggtggcag ggctttcctc agggccctta ctctgacct      60

```

ggacgagggc	ggggcttctc	caaggaggct	ctctgactgc	caccctgccc	tgctgcccc	120
gccctgcaca	acatgcagcc	ctccggcctc	gaggggtccc	gcacgtttgg	tcgggtggcct	180
ctgctgagtc	tgctgctcct	gctgctgctg	ctccagcctg	taacctgtgc	ctacaccacg	240
ccaggccccc	ccagagccct	caccacgctg	ggcgccccca	gagccacac	catgccgggc	300
acctacgctc	cctcgaccac	actcagtagt	cccagcacc	agggcctgca	agagcaggca	360
cgggcccctga	tgcgggactt	cccgtcctg	gacggccaca	acgacctgcc	cctggtccta	420
aggcaggttt	accagaaagg	gctacaggat	gttaacctgc	gcaatttcag	ctacggccag	480
accagcctgg	acaggcttag	agatggcctc	gtggggcgccc	agttctggtc	agcctatgtg	540
ccatgccaga	cccaggaccg	ggatgccctg	cgccctcacc	tgaggcagat	tgacctcata	600
cgccgcatgt	gtgcctccta	ttctgagctg	gagcttgtga	cctcggttaa	agctctgaac	660
gacactcaga	aattggcctg	cctcatcggt	gtagaggggtg	gccactcgct	ggacaatagc	720
ctctccatct	tacgtacctt	ctacatgctg	ggagtgcgct	acctgacgct	caccacacc	780
tgcaacacac	cctgggcaga	gagctccgct	aaggcgctcc	actccttcta	caacaacatc	840
agcgggctga	ctgacttttg	tgagaagggtg	gtggcagaaa	tgaaccgcct	gggcatgatg	900
gtagacttat	cccatgtctc	agatgctgtg	gcacggcggg	ccctggaagt	gtcacaggca	960
cctgtgatct	tctcccactc	ggctgcccgg	ggtgtgtgca	acagtgcctg	gaatgttctc	1020
gatgacatcc	tgacagcttct	gaagaagaac	ggtggcgctg	tgatgggtgc	tttgtccatg	1080
ggagtaatac	agtgaacccc	atcagccaat	gtgtccactg	tggcagatca	cttcgaccac	1140
atcaaggctg	tcattggatc	caagttcatc	gggattgggtg	gagattatga	tgggggccggc	1200
aaatccccct	aggggctgga	agacgtgtcc	acatacccag	tcctgataga	ggagttgctg	1260
agtcgtggct	ggagtggaga	agagcttcag	ggtgtccttc	gtggaaacct	gctgcggtgc	1320
ttcagacaag	tggaaaagggt	acaggaagaa	aacaaatggc	aaagcccctt	ggaggacaag	1380
ttcccgatg	agcagctgag	cagttcctgc	cactccgacc	tctcacgtct	gcgtcagaga	1440
cagagtctga	cttcaggcca	ggaactcact	gagattccca	tacactggac	agccaagtta	1500
ccagccaagt	ggtcagttct	agagtcctcc	ccccacatgg	ccccagtcct	tgcatgtgtg	1560
gccaccttcc	cagtccttat	tctgtggctc	tgatgaccca	gttagtcctg	ccagatgtca	1620
ctgtagcaag	ccacagacac	cccacaaagt	tcccctgttt	gcaggcaca	atatttcctg	1680
aaataaatgt	tttggacata	gaaaaaaaaa	aaaaaaaaag	ggcgccgct	ctagaggatc	1740
cctcgagggg	cccaagctta	cgctgcatg	cgacgtcata	gctctctccc	tatagttagt	1800
cgtattataa	gctaggcact	ggccgtcggt	ttacaacgtc	gtgactggga	gatctgtag	1860
cttgggatct	ttgtgaagga	accttacttc	tgtggtgtga	cataattgga	caactacct	1920
acagagatgt	aaagctctaa	ggtaaatata	aaatttttaa	gtgtataatg	tgttaaacta	1980
gctgcatatg	cttgcctgct	gagagttttg	cttactgagt	atgatttatg	aaaatattat	2040
acacaggagc	tagtgattct	aattgtttgt	gtattttaga	ttcacagtcc	caaggctcat	2100
ttcaggcccc	tcagtcctca	cagtcctgtc	atgatcataa	tcagccatac	cacatttgta	2160
gaggttttac	ttgctttaaa	aaacctycca	cacctcccc	tgaacctgaa	acataaaaatg	2220
aatgcaattg	tggttggtta	cttggttaat	ggagcttata	atggtaccaa	taaagcantg	2280
catcacaan	ttcccaaata	aagcattttt	tccctggaatt	taaatggggg	ttgg	2334

&lt;210&gt; 16

&lt;211&gt; 2608

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 16

ccacgcgtcc	gcacagggct	caggctgggc	tcagaggcca	ctgagatgcc	agctccttga	60
gagcagtggt	ggtgtcccag	gccagagcag	cctcttctc	tcctgggtccc	agaaaaaccc	120
ttgcagtata	tgggtggcctc	tgggtgggtta	cttctagcgc	aggcttcctt	ccttctctta	180
gctcctcccc	gtgccctggg	tgctgggtgc	tggatggatg	ggcgctcctc	agctcctccc	240
ggtgccctgg	gtgctgggtg	ctggatgggt	ggcgctcctc	tagctcctcc	cggtgccctg	300
ggtgctgggt	gctggatggg	tgggcccac	ggtgcaccct	tggtgggtg	cctgtgcccg	360
agtggcctct	gcagctctta	tgtctgcctc	taatgggatg	tgccgcctga	ctgcctcggt	420
cttaagggca	tagtgggtcg	gctaagatct	gatcgccagg	actgcgttct	gggcagggtg	480
tgggaaggcg	gaaccaagcc	gtccgtgccg	ctagggagcc	gagactgccg	gaaagaagag	540
cggcaggagg	ggcggtggtt	gtgcagcccc	acccccggga	ggggcttttag	gcactgggaa	600
ggaaagtcc	gttggaggaa	ttcgggtggc	gttcacaccc	gcctcgctgt	cttggagtct	660
tgatctgtcg	tcggcgggtc	gctggcacag	gactaaacat	ggctgagggt	gggctccagc	720
cagatctggc	tggggacagc	accgcgtggg	cccaggatcc	acccagagca	ggcaggcctt	780
gctggggccc	cagtcaagtc	cacttccagt	gaggagagcc	agccgggagg	tcagtgccag	840

8

agctctggtg	gagcccagac	cctgccttcc	ctgagggccg	cccctgtcgc	cgccctgggg	900
tccctgtcct	cctatcctga	ctcctgcccc	agggccacca	cccctgaact	gtgccctggg	960
gccccacccc	tccacctggc	cgactccatc	tctgggcctg	tcagtccacc	tgggtcctct	1020
ctgggccctg	atgcctggac	cctctgtgcc	aagcaccacc	aagcaaaggg	gatgaccttg	1080
ggcaccacca	aggtgtgag	actacagcca	gtgagccctt	gctggggggc	aaagtcattg	1140
aggggtgcctg	ggcccttcca	acctggaagg	aggaggggag	agagcaggca	gcagggcagg	1200
gggaagagga	ggagtgtccg	atctgcacag	agccctacgg	gcccagagag	cgccgcctgg	1260
ccctgtctgaa	ctgtagccac	ggcctgtgtg	tgggtgcctt	gcacaggctg	ctgggtctcg	1320
cctccagtgc	cgacctgggc	cgggtgcgct	gcccgtgtgt	ccgcagaaga	cgcccggtgt	1380
ggagtgggag	atctgcccgc	tacaggagga	gctgtctacg	gccgacgggc	cctcacgcca	1440
gccccgcaga	gaggcccctg	cctcctatca	ccgcaaccct	gggccctggg	gctccctgga	1500
gcaccgctac	cagctgcgct	tcctggcagg	gcccgtgggc	ggccggggct	gcctgccctt	1560
cctgccctgt	ccaccctgcc	tgggtgcccg	gctctggacc	ctgcgggagc	ggggaccctg	1620
tgcccgccgc	ctggcgctgc	tgagcctgct	ggcccttgag	cttctggggc	tgctgtctgt	1680
cttcacgccc	ctcctgtctg	tgggactgct	cttcgtgtct	ctggaccgct	ctggccgctg	1740
agcagagccc	aggacagccc	cgccgcaaca	ggccaggggg	cccagactgg	cccacgtccc	1800
catgcctggg	tgctgtgagg	cctgatgacc	aggctggaaa	aacccaaggt	tgggtccagg	1860
gcagtggcct	tcaatcaaga	cctcccattg	ctgaaccac	aaccagggct	accagaggc	1920
ctgaccctgc	agagtccatg	gctgcactgc	tgcccagaca	ctagctgaac	ccaaggacac	1980
cagcgcccaa	ggacagctcc	tggaggaggg	cagcccagca	ggaaagtctg	tgagcaggac	2040
cccattcacc	ctgcggcaga	cgggcaccgt	actggccacg	ggctgacgcc	ggccacactt	2100
cccctccgag	ggccagctga	gcacagcagg	catgaaagca	aacagagata	cagcagttag	2160
tcagttcctt	ggagagggca	gggactccgc	ccaccctgtg	ttcagataag	ggccagtgtg	2220
tgtccctgaa	ggtcaggcca	gccgggggag	gggtccatgc	tgcgaaaatt	cagcctgcaa	2280
aggctcctct	ccccacttga	tcaggcccag	accagggtgg	gggtggcgct	ggcctgtgtt	2340
gcaggggaca	agggcccacc	caggccttgg	aacataagct	ctgcccctgc	acaccctcat	2400
gtcaccacac	ctgggatgga	gacatcaggt	ggcccagcga	gagatggagg	actgatcctg	2460
gaacgtgaag	cagcttttcaa	taaaccagct	cctggggaaa	aaaaaaaaaa	aaaaaaaaaa	2520
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2580
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				2608

&lt;210&gt; 17

&lt;211&gt; 1291

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1279)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1286)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1290)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 17

aaacctcttc	tataggtaaa	gctggwacgc	ctgcaggtac	cgggtccggaa	ttcccgggtc	60
gcccacgcgt	ccggaaagag	gaaacataga	gggtccaaag	gaacaaagac	ataatgatgt	120
catccaagcc	aacaagccat	gctgaagtaa	atgaaaccat	acccaaccct	taccaccaa	180
gcagctttat	ggctcctgga	tttcaacagc	ctctgggttc	aatcaactta	gaaaaccaag	240
ctcagggtgc	tcagcgtgct	cagccctacg	gcacacatc	tccgggaatc	tttgtagca	300
gtcaaccggg	tcaaggaaat	atacaaatga	taaatccaag	tgtgggaaca	gcagtaatga	360
actttaaaga	agaagcaaa	gcactagggg	tgatccagat	catggttggg	ttgatgcaca	420

ttgggttttgg	aattgttttg	tgtttaatat	ccttctcttt	tagagaagta	ttaggttttg	480
cctctactgc	tgttattggt	ggatacccat	tctgggggtg	ccttctcttt	attatctctg	540
gctctctctc	tgtgtcagca	tccaaggagc	tttcccgttg	tctggtgaaa	ggcagcctgg	600
gaatgaacat	tgktagtctt	atcttggcct	tcattggagt	gattctgctg	ctgggtggata	660
tgtgcatcaa	tggggtagct	ggccaagact	actgggccgt	gctttctgga	aaaggcattt	720
cagccacgct	gatgatcttc	tccctcttgg	agttcttctg	agcttgtgcc	acagcccatt	780
ttgccaacca	agcaaacacc	acaaccaata	tgtctgtcct	ggttattcca	aatatgtatg	840
aaagcaaccc	tgtgacacca	gcgtcttctt	cagctcctcc	cagatgcaac	aactactcag	900
ctaattgcccc	taaaagaaaa	aggggtatca	gtctaattctc	atggagaaaa	actacttgca	960
aaaacttctt	aagaagatgt	cttttattgt	ctacaatgat	ttctagtctt	taaaaactgt	1020
gtttgagatt	tgtttttagg	ttgggtcgta	atgatggctg	tatctccctt	cactgtctct	1080
tcctacatta	ccactactac	atgctggcaa	aggtgaagga	tcagaggact	gaaaaatgat	1140
tctgcaactc	tcttaaagtt	agaaatgttt	ctgttcatat	tactttttcc	ttaataaaat	1200
gtcattagaa	acaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaggcg	gccgctctag	1260
aggatccaag	cttacgtang	cgtgcntgcn	a			1291

&lt;210&gt; 18

&lt;211&gt; 3129

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 18

cggcacgagg	ggagaccgaa	tggcgggggtg	aggggagcta	ctgcgtccgg	ctttgcaggc	60
cgttcctttc	aggggtgccc	cagcaatggc	acttcatect	ggcagcagcc	atttgctggt	120
tgcagtgcca	gtttcttggg	ttcttttttg	cattcctgga	atcagtttca	tactctcttc	180
ttggagttac	caagagtcgc	cagtcagttt	cctctccgtg	gaaggctgag	tccaatttcc	240
atagggtctt	gtgccagact	tttagtctct	gataactccg	aaatccttcc	ttctgtttcc	300
taagctctac	aggtattcta	gtttttctcg	caatttttca	ttactaggct	acatcagtgt	360
cttttttctt	tctttttttt	tctcactttt	tatttgtacc	tttagtcttc	caacaattaa	420
caattcccca	tgttaaatcc	tctctgttta	aacaactaac	gggctctctg	tattccaact	480
agatctcggc	tgatgcaaat	gccgtagaca	atcaatcgat	caatcaatca	gcaaatccaa	540
tttccctcta	gccactaaag	ggacttaata	taattgaatg	attataaaat	tctgatccca	600
tttatctcaa	gccattgaaa	atatttttctg	tgtgattata	agatacgtaa	ttttaaaaaa	660
tgattttcta	aatcaggatt	cttaccagat	tatttcacag	cttattcact	cattgattgc	720
tttccctcca	cccttggttag	cttttgtatg	cctgtcctgc	attttaaacc	aagggtatgag	780
gaggataaag	taaatagtaa	atgggtatgat	agtggccaaa	actttgaatc	tttttcccta	840
ttcatcatct	ttgtgcaagt	tggtagggag	gtccactaat	tcagcctccc	cattccctgt	900
cccttcaaag	taaatcagaa	gcttcaggta	aaaagcagag	aagctaccca	agttcttttc	960
atttattttta	tcatcatact	aataattttg	ctttctgaaa	tgattactct	tacaagcagc	1020
aaatatatatt	tttaggaagt	tagaattatt	gatgtcacgt	tttcatattc	agtatcccaa	1080
gtaaatatca	ggtaaattca	ggaaaataaa	tacaaccatt	ttcttggtta	attgtttttta	1140
aatttaatat	atgttcattt	attctttttaa	atataatatt	cataaatggg	gactatgtat	1200
tccatgacat	aaaaactaag	tatttttagt	gggaaattct	ttgtggaatc	actgaaatga	1260
taactagtgt	tagcaatact	tctgtttcat	gagacgagtc	actttatgaa	gattgcaaat	1320
tattttttgt	cccctctaga	gttcattttca	tgggaaattt	aaatttttatt	gtcttaactt	1380
taatgttaaa	catttccttc	acatatggta	tggccagaca	tgttttcttc	tttgctgcta	1440
ttttacattt	tagatgtagt	ccattcagga	ctctgtattta	catgctaata	aagacaactg	1500
ttgaagtgat	gcaacacttt	cactctttat	tcttgattta	tgaggaaatt	ctgtggcctg	1560
atagacaata	tgatatactt	tggccctaga	agatccttaa	tgaggaaatt	ctgtggcctg	1620
ttttctcttc	tgtgctatatt	tttgtcttct	aaggcaggca	tttttatcag	ttcagtttgt	1680
tgctgtgact	agttcaaact	tgttcagggc	tttatgctgt	aaataagtaa	tcaattagcc	1740
tatatatata	aatatataac	caaccattac	aattatctaa	gaatgtcagg	gagttgtgaa	1800
gtggaaacaa	attaaagagt	tgcaaaatta	aaagagctcg	tcttccaaca	gaagtgtgct	1860
gacatatatt	aaagaataaa	tgaattaaag	tcaccaaaaga	ggaacaggag	ttaaaaagaa	1920
gaccattgtt	tagatggagt	caggctgggt	cactgccaaag	gaaagcaaac	aggattttctg	1980
aaggcttggg	atgaatttca	gcttgcacat	ttccgctgga	gggggtgatt	tccccaatat	2040
catcttagat	ccatagcttg	ttcttgacag	agtggcagta	cctttcctcc	actgcactct	2100
caccactagc	atagatgtaa	aacacagtaa	gtactcagaa	actacttgaa	gagtgcagtt	2160
atcagtagag	atgatcgaaa	catttgtttt	tctagggaaat	atttttgcct	ttcttcttcc	2220



10

agaatcctct	ggttataatg	tgctcactgc	taggtcacca	gtcataaaac	attatgtaga	2280
ggttactggt	cattatccta	atatatttat	caaaaaatct	ggagtatatg	aaactgcctt	2340
tcattgtaac	attagacaaa	aacatttatt	ctatcaaata	cagacttaaa	actgccacca	2400
aattggaaga	atatgatctt	aaatttaaaa	aaaccccata	tacttaaaaca	caataatcta	2460
tctttatcta	ccttcctaaa	cattaatgca	tgagacttct	cttaatatca	tgaatatgca	2520
ctatgatttt	tatgttacat	ctttttcttg	tttccattta	tgctagtga	atttattagt	2580
atccttcatt	gaaacactat	tattttccat	ggaggaaaat	attttatttc	ctttatgaaa	2640
aggggtgctt	acttctgaaa	tacagaattc	attttgtttc	ggattctggt	tgtttgtttt	2700
cacatcaact	tcattctaaa	tgttcattca	aaattatttt	tcagacttga	gttcagggtga	2760
aaatgttaat	ggaattaaaa	agtagttagt	caaatagaat	ttcaatatat	aagtcaaatt	2820
tgaagaaaac	tgaattaata	aaggattcta	agtttataaa	gaaatcaaaa	ttatgctttc	2880
aaaaatatta	tgtaactgga	gaaagtaatt	ttattttgag	atctctcatg	attccttaatt	2940
atatattttt	tcttattagt	atgcaacctt	ttggtaacat	atataggaga	cagttaattt	3000
tggtgaagact	acataattca	tacattcttc	atcttgacat	acatgatcaa	atactttata	3060
acgtccacta	gccttgctct	tgtccactgc	aatggaaaaa	aaaaaaaaaa	aactcgaggg	3120
ggggcccg						3129

&lt;210&gt; 19

&lt;211&gt; 3629

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 19

ccacgcgtcc	gccagaggct	gggaaggaaa	aattattccc	atcttaaaat	atcttaaaaa	60
tatgtacaga	tttattgtaa	atcaaaat	tcctcataac	tacagcccct	tcccacacat	120
aaaattcccc	ctaaatcagg	aacataatgt	acagcctggt	tttaacttgt	atttttccat	180
tcacactatg	tcacaaaaaa	attctcatgg	tgatccatga	tttcacaggt	cctgtccatg	240
tgtttcctga	gaagacagtc	ctggaatgga	attactaatt	caaaaggcat	ttttaaggct	300
gtttacacata	ctgccaaatc	ggccttcagg	aggatgacac	aaattccacc	tatgtcatct	360
gagcgtgccca	ctatctgctg	ctattgcccc	atgtttgagg	ttttgtcctt	tctcctcatt	420
ctcatgggca	actgtcaaca	tcaccaccac	tgtgatcctg	aaaagaaagg	ctcagaagac	480
atagacagtc	tgggattttt	ttttttccct	acctttcttt	gcaactgtagg	gaaggagtga	540
gatgtggctg	caggattggg	gaatagctaa	gtcagaaatg	agcagcagtg	caactgtggc	600
aatcttggag	gggggtctca	gccatgaaag	aaaacagggg	gtgccttgcc	agctggaggc	660
tgcagaatct	tagcttgggt	gttttgccct	ctccgtggaa	ctgacagatg	acaggaatga	720
cgtggagtat	tctcatcaag	cgtctcagca	tagtcagtc	ctgagatggc	aaactttagg	780
caaagacacc	tgggctgaaa	ggtcccacat	ttaaaacttg	tcagtcagat	aactgcccac	840
caccacccct	tgttaagatac	ctagatatgg	gaaccccaca	ctcttcccac	ataagactct	900
tgtatgactg	aaagactttg	tggttaaaaa	catatcttaa	tctaacctgt	attttccctg	960
ctgcaatttc	agcctttttc	tatcttttca	tgggtgaatga	acaccacaca	ggctttcaga	1020
tcagacaact	tctggctccg	tcatttacta	gctgtatcag	ttaacaccag	ctgctgactt	1080
cacttctgca	tatgcattcc	aaaatttagtg	tcttaaaaata	acaactactt	attacttttc	1140
acaagcctaa	aagttggctg	aatgggtctg	ctgatctgac	ttgggttaatc	ttggctggac	1200
actgtctgat	tcagcatagg	ttgactgtcc	tctgcctcat	gtctcttatc	atccataaat	1260
tgcccaggct	tgcttgccctg	aaggcagaag	agttcccact	agcaggagag	ggcaagcact	1320
tttcaatgca	caagcacctt	tcaagcctct	atttgcatca	agtttgctaa	tgccccgctg	1380
accaaagcaa	atcatgtggc	ttacttcgga	gtcagtgagg	gtaggcatta	cccaaagaaa	1440
gtggctatag	ggaagcatga	agatctagag	ccatttagtac	aatcaatcta	ccacattaac	1500
tttgtgacct	tgggcaattc	ccttagcctc	tctgagcttt	aactctataa	tctataaaaag	1560
tgcagaat	ttttgggatt	ctttgtctata	atgtatgtaa	aatgtctgct	acacaataaaa	1620
catgccatat	ttattaat	tcttttctcc	tccttttcat	catttgaatt	ttttcttaaa	1680
tgcatatata	tatgtgcata	tatatatgct	catctcgata	tatatatatg	tatatatata	1740
tatatatatg	acttactctg	tcattccagg	gagagtgcag	tggcatgatc	acagctcact	1800
gcggcctcga	cctcccaggc	tcgagtgatc	gtcccacctc	agcctcccag	gtagctggaa	1860
ctacagatgt	gtgccaccac	acccagctaa	atttttttgt	atattttata	gagaagggtt	1920
tcattatgtt	gctgaggctg	gtcttttaact	cctgggttca	agcaatccac	ccaccttgac	1980
ctcgcaag	attgggatta	caggcatgag	ccaccatgcc	tggcccttaa	atgcatcatt	2040
cttataacatt	cctttataca	tacaaagctc	gaagataata	ataataaaca	ttattcccac	2100
ggcttctgtt	gaatcccca	ctcctagcct	gtgcctaata	aagtgtgccc	tcaattcatt	2160

11

ccaatggtct	ggaatttgaa	gactaaagga	atcaaaggg	atctcttccc	tttgttctg	2220
ggaccatcaa	gttgtgaaat	gacctcttgg	ctttctgac	tgggtccact	accacaaatg	2280
aaatagaaag	gagttaacac	atctgcatta	actgatgatt	cttaatagag	gcaataaaat	2340
gagaaacccc	catttcctaa	tttaaatgtg	gctttgggga	taggataagt	atcctggccg	2400
ctccttagct	aatcaggctc	cagaaggac	attaattgca	gccagcgcat	cctgctgggg	2460
agacccaagg	ctgccagctc	ccttgctgag	aatggaatag	aattttaatt	ggctgtgaaa	2520
gcatcatggt	aggaggtaca	gggaagagtt	ggtctgctcc	cccaaagcag	gttagggaag	2580
ctccaaggga	aaggactcaa	acaggtgcat	ggcagcaaga	acctcccatg	ctcaggggccc	2640
ttcccagctg	caacaccaca	gttccaattc	cagggctgga	ctccagccta	ggtgaaggac	2700
cttctcacac	ttgggcgatt	agtactggag	cacggagcct	gtgaatcttc	tcatctctgt	2760
gcttccccc	acttccatga	gagaagcacc	acatagttac	tgctacttgt	attgtcaatc	2820
ttcgacttca	gagaggtctc	caaatataag	ctcaactccc	cagtgtactg	tgtgtgtgcc	2880
attagtcct	gtgtatggtc	tgaagcaatg	gcacagcatt	cctgtactgt	tgtttcagag	2940
caatgtgaat	ttatttctact	cacaataaat	taatctaact	atgcttaggg	caaagttgtt	3000
ctcatggtaa	tgagtatgtc	ctttctgcca	tggcaatttc	tggaaacacac	atatagacaa	3060
gaatgacatg	tgaaggtcaa	tagatgagac	tatgaaaaac	aagacaatat	agcttttttag	3120
cataaatgta	caatgatgca	tgtgggtttt	ggagattggt	ggacaaatta	cttaacctct	3180
tgggtgcctta	gttttctaat	tattaaaatg	agggcaacag	tattgctgat	ctaataaggc	3240
ttttggaaag	aataatgaga	tgtaaagagc	ttagaatagt	gcagagtctt	aataaacatg	3300
tcttcgtgat	tccctctagt	gaagtgcac	agagaagtgg	gaccagtcac	tgacccaccc	3360
aagctctgtc	aacctagtgg	tatagtaaat	ctgttggcaa	aaagcaaagc	agaacaacaa	3420
aaaaacgcaa	gcaaatcat	tgctttcagg	aaactatttt	gtgagacaga	agcttgataa	3480
aaattcaaaa	tttgcagtaa	actctacttc	attaatactt	taaaactcac	agcaataaac	3540
cattacttag	aacaagtaag	cagttttgtt	tttattgtgc	atcttgatat	ttattttcaa	3600
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa				3629

&lt;210&gt; 20

&lt;211&gt; 1144

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 20

tcgaccacg	cgccgcgct	tgacttgctt	ccagacaaag	gttgtctcaa	gtttgttctg	60
caaaccgagt	tctggagaac	gccatcagct	cgctgcttaa	aattaaacca	caggttccat	120
tatgggtcga	cttgatggga	aagtcacat	cctgacggcc	gctgctcagg	ggattggcca	180
agcagctgcc	ttagcttttg	caagagaagg	tgccaaagtc	atagccacag	acattaatga	240
gtccaaactt	caggaactgg	aaaagtaccc	gggtattcaa	actcgtgtcc	ttgatgtcac	300
aaagaagaaa	caaattgatc	agtttgccaa	tgaagttgag	agacttgatg	ttctctttaa	360
tgttgctggt	tttgtccatc	atggaactgt	cctggattgt	gaggagaaag	actgggactt	420
ctcgatgaat	ctcaatgtgc	gcaatgtacc	tgatgatcaa	ggcattcctt	cctaaaatgc	480
ttgtcagaa	atctggcaat	attatcaaca	tgtcttctgt	ggcttccagc	gtcaaaggag	540
ttgtgaacag	atgtgtgtac	agcacaacca	aggcagccgt	gattggcctc	acaaaatctg	600
tggctgcaga	tttcatccag	cagggcatca	ggtgcaactg	tgtgtgcccc	ggaacagttg	660
atacgccatc	tctacaagaa	agaatacaag	ccagaggaaa	tcctgaagag	gcacggaatg	720
atctcctgaa	gagacaaaag	acgggaagat	tcgcaactgc	agaagaaata	gccatgctct	780
gcgtgtattt	ggcttctgat	gaatctgctt	atgtaactgg	taaccctgtc	atcattgatg	840
gaggctggag	cttgtgattt	taggatctcc	atgggtggga	ggaaggcagg	cccttcctat	900
ccacagtga	cctgggttacg	aagaaaactc	accaatcatc	tccttcctgt	taatcacatg	960
ttaatgaaaa	taagctcttt	ttaatgatgt	cactgtttgc	aagagtctga	ttctttaagt	1020
atattaatct	ctttgtaatc	tcttctgaaa	tcattgtaaa	gaaataaaaa	tattgaactc	1080
atagcaggaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	1140
cgag						1144

&lt;210&gt; 21

&lt;211&gt; 1443

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

12

ggaaccattg	gcctatat	ggttggatct	attattatga	gtgttgttgt	ttttgtgcca	60
ggaacattg	tagggaagta	tggaaacacga	atttgccttg	cttttttctt	aagcatacca	120
tatacttgtc	ttcctgtctg	ggctgggttc	agaatctata	atcagccatc	agaaaattat	180
aattaccct	caaagggtat	tcaagaagcc	caagcgaaag	acctgctgag	aagaccattt	240
gatttaattg	tgggtgtgtg	tctcctcctg	gcaactggat	tttgcctggt	cagagggttg	300
attgctttgg	attgcccatc	tgagctctgc	cgattatata	cgcaatttca	agagccctat	360
ctaaaggatc	ctgctgctta	tcctaaaatt	cagatgctgg	catatatgtt	ctattctggt	420
ccttactttg	tgactgcact	gtatggctta	gtggttcctg	gatgttcctg	gatgcctgac	480
atcacattga	tacatgctgg	aggtctggct	caggctcagt	tttctcacat	tgggtgcatct	540
cttcattgcta	gaactgctta	tgtctacaga	gtccctgaag	aagcaaaaat	ccttttttta	600
gcattaaaca	tagcatatgg	agttcttcct	cagctccttg	cctatcggtg	tatctacaaa	660
ccagagttct	tcataaaaac	aaaggcagaa	gaaaaagtgg	aataaaaata	ttacttcatg	720
ttcctccttt	ctaaattact	aacttttggt	atactggtac	tgatattttg	tcccatttca	780
ctctcttctc	atacgtgagt	acttaagaat	atgtacattc	ttgctctgca	ctgtatgtgt	840
gagctatatg	gtatttgtga	aatttttttt	gaaggaaaat	ggaaattcct	gagaaacagt	900
ttgtttaaag	aaatatattc	aaaatcattt	gtgaataaac	ttgatcatcc	atctcaatat	960
tgtttgacat	ataaaataat	tataagtgtg	aaaaattaca	atttagtgcc	aacagtagtg	1020
agcatgaaat	gaaactattc	aaaagagaat	atggcctgtg	catattaaaa	aattcaaaaac	1080
agtgaatgca	gactggagga	gtaacttttg	caaataagat	gaatatgctt	cattattaaa	1140
ctcaatataa	aaggcaaatac	atcagaatat	ttttaaatgt	tgtttgaaaa	atgttttccc	1200
aaggaaagtt	tattatttgc	tgctgtttca	agaaaattac	tttactaaa	tttttttgtg	1260
tgaatttaaa	cagctaaata	gggatcagta	actttatctc	tatccttaat	gaacatttgt	1320
tttattgggtg	gctggaaata	tttctattgt	atttctgtgt	atatttttta	taaaattatt	1380
tttggcctct	taaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaactc	1440
gag						1443

&lt;210&gt; 22

&lt;211&gt; 1053

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

cttcgctggg	gggaagaagc	cgagatggcg	gcagccagcg	ctggggcaac	ccggctgctc	60
ctgctcttgc	tgatggcggg	agcagcgccc	agtcgagccc	ggggcagcgg	ctgccggggc	120
gggactgggtg	cgcgaggggc	tggggcgga	ggtcgagagg	gcgagkccc	tgtgtcctca	180
gccatcccaa	gaagggtttg	ctgggtccctc	ctttccccc	gtcccacgag	gccacctggg	240
ccagccctt	gtcctctgcc	ttctgctggc	agaggagcag	ctggactggg	gcctttggca	300
cagcagccgg	tgtctcctgc	gcccgcctcc	cccattggccc	catgcagccc	caggggcttc	360
ccccctgccc	atggagtaga	gcccagatc	ctggcacta	tgccagttct	gacctcgcat	420
ccccctaccc	cgagcccatg	cagtctggga	acatgccgc	ttctctccag	cctctgtgcc	480
tttgttccag	gtgggtctcac	cctcctgtcc	ctggctgggc	taggtgggtc	tgtccaggct	540
cctgcagcgc	ccccctcact	ttgacactgg	actaggatgc	agcctccctt	ctgtgtcccc	600
ttgaggggtac	cctgggtccc	ctcatcaggg	gcagaggcat	gaaagagtcg	gggctggatg	660
gccaggggct	tctgggccc	acgcctagt	cagccctgg	ggcgtgggtt	tgacatttgt	720
ctgcctgggtg	caaacaagga	atccttgcct	ttaaggtgac	aggccctcca	caggcttcca	780
gacttgaagg	aaaaggttta	agaaagaaaa	caaaaccaac	agtttagtggc	cggtgagggc	840
ccaggctggg	cagcgtcccc	tcttgcacac	ccaggggcct	ccctttctgc	tggagtcccc	900
tgtgtcctcc	accaccccc	gccgcccagc	atcctacctg	gactgcgggtg	ctacgagggc	960
ctgcgggcct	ttgctgtgtg	ccaccctccc	tgtaagtcta	tttaaaaaca	tcgacgatac	1020
attgaaaaaa	aaaaaaaaaa	aaagggcggc	cac			1053

&lt;210&gt; 23

&lt;211&gt; 741

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 23

gtagacttct	ggaccttgra	tgtcactgag	tttaagagaa	aggtagtccc	aggactctgg	60
------------	------------	------------	------------	------------	------------	----

13

tctagcagct	gagtaaggct	caggccttag	aatggcttag	cttgcctccat	tgcaatgcat	120
catcctggat	gttaaaatcc	agctgtctct	ctgaaamcta	aatatgaaag	actgagattt	180
agtcaacttt	gctgagattt	aatttacata	taataaaaatg	aatcatttta	actgtgcagt	240
tcagtaagtt	ttgaaaaatg	tctatacaga	ttcatgtaac	tgccaccaa	attgagaaag	300
gacacttcca	ccatctcaaa	agattctgtg	tggtcctttg	tagtaagtct	cttctacccc	360
atccctagac	aacagctgat	gtgctgtcat	tgtacacata	tattagcttt	gcctgtccta	420
gaacttcatg	ttaatgggaa	gcatcctgta	tgtactgttt	tgtgtctggc	ttcttcagtg	480
tattttttgaa	atztatccac	attgttgtgt	gtatccaaag	tgtgttcttt	ttcattgcca	540
aataatgttc	tgctatatga	atatattaca	aaatatttgt	ctgtttatct	attggtggat	600
ttttgcattc	gttccagttt	ggggctatta	tgaataaagc	ttctgtgaaa	attcaaaaaa	660
aaaaaaaaaa	aaaatgacc	tcgagggggg	gcccggwacc	caaaacggag	tatttccctt	720
tttcccccc	cccgcccc	g				741

&lt;210&gt; 24

&lt;211&gt; 946

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 24

ggcagcgagcg	aaagcctctc	tcttaacaac	ggtgccgcac	agctttgccc	ttgaaagcat	60
ctctactgga	cgggaacaca	ctcatgtgcc	ccgctccctg	accagccaa	ggctgccctt	120
tcctctccaa	ggctgagatg	ttgccggtgg	tcccatgaga	gcctgccc	gggctcaggt	180
gcccctttac	cttctgctgg	atggacatct	ggctgtgagc	caggctgggg	tcctggccgg	240
ggtgagcgga	ggcaggggtg	gacggaggct	tcgagggccc	atcactagta	gggtcattac	300
ctcttgccaa	cagccggggg	tgggagtctg	ggtctcgctc	aggccagagc	ttctcaacct	360
ggagtccctg	ggggtggctg	ccaaagggtg	gtatgacaag	cacgtatccc	tggaattttc	420
cggggagagg	tctggggctt	tggtcacatt	ctccaagggc	tgctgggctt	cggagcagtc	480
cccccccatg	tctcagccac	tacagggtcc	ctctctctcc	ttgcacccca	gaccctccgc	540
tgccctggta	atgagcagaa	ggaaaagtctt	ggggtgtgtc	caaagtcagg	agagcaaaat	600
atgccaggca	aaagctccc	ggaaaagccg	gaggagtctg	gggtggccac	cgggatgtgg	660
agcagcgagg	gcaaagacgg	tgaacacagc	cctccagctg	tctgagcctc	agttttctaa	720
tctgtagaat	gggatgatc	atacctgcct	cacaagaatg	ttgagacaat	tcacagagac	780
gttctggagc	ccctttcccc	cgagaccggc	attcatgagt	ctgctgggac	cagaaaaccc	840
atctcagggg	cccagcgggg	cacccaggag	agtctggcgg	tgcaagcgct	gtataaacca	900
caagcgttct	ccaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaa		946

&lt;210&gt; 25

&lt;211&gt; 831

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (5)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (10)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (11)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

14

&lt;221&gt; SITE

&lt;222&gt; (15)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (27)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 25

catcnacggn	naccnctact	ataggtnaag	ctggtacgcc	tgcaggtacc	ggtccggaat	60
tcccgggtcg	accacgcgt	ccgggggaaw	tcccagtcga	tttttccaag	cagtactccg	120
cttcctggat	gtgtttgtct	ctcttggtcg	cactggcctg	ctctgctgga	gacacatggg	180
cttcagaagt	tggcccagtt	ctgagtaaaa	gtttccaag	actgataaca	acctcggaga	240
aagttccagt	tggtagcaat	ggaggagtta	cagtgggtgg	ccttgtctcc	agtctccttg	300
gtggtagctt	tgtgggcatt	gcatacttcc	tcacacagct	gatttttgtg	aatgatttag	360
acatttctgc	cccgcagtgg	ccaattattg	catttggtgg	tttagctgga	ttactaggat	420
caattgtgga	ctcatactta	ggggctacaa	tgcagtatac	tgggttgga	gaaagcactg	480
gcatggtggt	caacagccca	acaaataakg	caaggcacat	agcagggaaa	cccatctctg	540
ataacaacgc	agtgaatctg	ttttcttctg	ttcttattgc	cctcttgctc	ccaactgctg	600
cttggggttt	ttggcccagg	gggtgaactt	tatttcattt	ccmcagggtg	aaactgaatg	660
ggcagttcat	gktaaaatcm	cttttcatgg	aaagagctct	atgtaacagc	ataataaaac	720
tgsctaccta	gcagcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaagggcg	780
gccgctctag	aggatccaag	cttacgtacg	cgtgcatgcy	acatcatagc	t	831

&lt;210&gt; 26

&lt;211&gt; 1294

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

ctgcagaatt	cggcacgagg	ttatttcacc	tctcttgccc	tcagtttctc	tgtgaaatca	60
ggaatttaaca	tggctctctga	gaccccttc	tgatggtgaa	tgtgtggttt	ggtgattttg	120
tggccctgca	tcatgacctt	atttagttct	ctttcaacag	gggatgtttt	actgccttgt	180
aaaatcctcg	tgggactgcy	tgtctttata	ggagccaggg	tgtaaatgaa	cagaattcag	240
attggttcta	atatatttta	cctctaaaag	aaagggcatg	gggaggccat	gaccttaaa	300
caggtttttt	ctgttgtctg	tgaagcctgt	gatgattgag	agtggctggg	actggcggga	360
cgatgtttgg	gtggaagagg	gaggccatct	cgatgcgccc	cgtcccgggg	aggcaccag	420
cctgtaagga	ggtgatgtct	atctacactg	agcgcaagga	ccctgaaccg	ggggaggctg	480
aggcggggcc	tcttgattcc	caccctgtcc	cccagtggct	aggctagtgt	ggcccgggaa	540
atgacttcca	tctctccctc	caggcatatt	taataagagg	ccagtatttt	cagattctgc	600
cgcttctgga	cgaatgtctc	agagagctgg	gaggcgccct	ggaggatgga	acccttccct	660
gagcgttgtt	gaggtgtgtc	gggggtgccg	tggcacaggc	cccctcccct	ggggggcatc	720
actgttccct	tgtcttgcac	ccccgctgtt	tcccctgccc	ctgaacaggc	gtggagatgt	780
gcacggggaca	ctcggaggcc	ggatgctcaa	cagagtggag	tgccgcgacg	gtgtggccgc	840
agcctggctg	tgccttcacg	acgcagctgc	aatcagagga	gctgtgggac	gctgtcccat	900
gtggacacag	ccactcact	gggtgctgct	cctgtgctgg	gcgctgcact	tttattgtcg	960
ttaaaaattt	atattaagat	gcggccgggc	atggtggctc	atgcttgtaa	tcccagcacc	1020
ttgggaggcc	gagacgggcy	gatcacgagg	tcaggagatc	gagaccatcc	tggcttacat	1080
ggtgaaaccc	cgtctctact	aaaaatacaa	aaaaattagc	cgggtgtagt	ggtgagtgcc	1140
tgtagtccca	gctactcggg	aggctgaggc	aggagaatgg	cgtgaaccgc	ggaggcggag	1200
cttgcaagtga	gccgagatcg	tgccactgct	ctccagcctg	ggcgactgag	cgaaactccg	1260
tctcaaaaaa	aaaaaaaaaa	aaaaaaaaact	cgag			1294

&lt;210&gt; 27

&lt;211&gt; 1656

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

ggcacgaggt	tcacagcacc	tgatttgcaa	ggcagctata	caagttcctg	gactcttgta	60
gttccggagt	gtttcacctg	accttaagcc	caccccatcc	atctttaatc	aagaaaccat	120
gtgctttccc	gcatacctgt	gttccccctt	cacgtgtctg	ctgtctgtgt	ggaagcctgg	180
cctggcgcat	gctgtggtgc	actgcatgct	ggaacccgtg	gagtttgac	gcgtggtaca	240
gtatgaggcg	ggtcacgttt	tgtagtgtgt	gccgtgggct	cccagaaaac	aagttaaagt	300
gtgtgctgaa	atagatttta	ttgacataaa	ataagcctta	ttgctaaatt	taagagaatg	360
tgttacaagt	gttttttgct	aaacatcagt	attgattatt	ctacatgatg	tacttattga	420
cataacaacc	tgaattcttt	gatttttagac	aattttctct	caagttgatt	cagctgcatg	480
actctcagaa	atcagtcatt	ttttattgta	gattgctggg	tttcttctct	tagtttgat	540
cgtgtatttt	cctcctgtgg	agaaaatgtg	gttggcaaga	aatgccatat	tttaaagctg	600
tatcgtggct	gttaatgcag	aaaacaccag	tgtactgcag	gctgtttggc	agtggggctg	660
gggctgagtg	tcctgccctc	agtggcctgt	gtctgtgctc	ttgttcgctg	acatgcagat	720
acagggggcag	atctgagggg	ttgatggagt	gcagaaggcc	acacgtgtgg	ctttctgtaa	780
atgcagaaac	atggaatcct	tgagcagaca	cttgtcttct	ggagcacctt	gcatggattt	840
cgctcctcta	tgtctcattg	ccgttaatag	agtgggtggg	gttgtgttat	gagaaatttt	900
gtctaaccctg	gcttctgaaa	tttctcaaac	taaataattca	tgctgttttg	tgtttttctt	960
aatgactgag	gctagtata	ttactcagaa	aagtaacagt	aacttgggtc	ttctgagcgt	1020
caggatgttc	accatttaac	ttgtttctcg	ttagtgtcta	gtacgtcggc	tttcggtagt	1080
gtagggtgtg	gttctgtgtc	ctttcccgtg	tgtgcctgca	ctagtggcag	cctctgcttc	1140
caggctcagtt	tagagtagac	tggtctgtgt	attgctagca	agtagttgct	gttaccagat	1200
gtagccatga	agcccagctc	cttggatctt	gacatatatg	ttccaggcaa	agtacgtaat	1260
ccagacgttt	ctaactcttt	ctagatgatt	gcaattgttc	tccatgttgt	ctgttaggcg	1320
ttatgttaat	tctcgatcta	acagtgtgcc	tgtaacatat	atggtagtga	agagacatca	1380
catgcagaga	ccgttttctt	tttatcaact	acaggtccgc	tatcgacgag	agcaccttct	1440
tgctaggcag	tcaccctact	tcccgttgtt	ggaggatttg	atgagagacg	gcagtgatgg	1500
tgctgtcttc	ttagctgtga	ttcactatta	ttgccagag	cagatgaaac	tgatgggtga	1560
gtggagaatg	cttccctgaa	cagatccgaa	aaggcttaaa	ggaaaattat	agtgtacatt	1620
gatccacata	tatatataaa	aaaaaaaaaa	aaaaaa			1656

&lt;210&gt; 28

&lt;211&gt; 1350

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

gtgacttcta	tattcaatag	atttttgtaa	atgttaaaac	atctatattt	aaatgttaaa	60
acactaaata	tagagagggg	ctttatttca	atcatagagc	aacaacaaaa	ataatgctta	120
tagctaaact	gcctgttcta	gaaagcatct	gctttttcat	gttattccta	aatcctcttg	180
tcatactttt	gtcattgaac	aatgctctcc	ctctcgtctt	ccatcctcat	tcagaatttt	240
tagaagacca	caatcgtgga	gatacactac	ccagtattgt	ttgatacatt	tttatttgat	300
aaacattcag	tgcaggaaac	tgtgatttgc	tatatgttta	tgtatataat	cttattctgt	360
agtcatcaga	atgttaatgt	aaggtacatt	tgatttttat	tttttacatg	tgtagttttc	420
tttcttcaca	gtcaaagcat	ttatattatt	gggggtgggg	gcagggaatt	aagttgggtg	480
gctcgaaaat	ccattcatat	gtatctgtct	acaaatgtct	ggggataatt	taaatttgaa	540
acctaagtta	tatatagtgt	ggcaatgctc	ttcttcaata	tttacaataa	taggatgatc	600
tacaagaaaa	taagtttctt	tttgcaaat	tttatcatac	taaagttgtt	cttttaattt	660
agcatatcta	aaataggatt	tagttcagtt	tagctcacac	aggtgtttgc	tgacattcat	720
tggccattta	atacagtgtt	gagtgttctt	ccgtaaaagt	ataagtgtta	acactacgaa	780
gaaatgcaca	cgatcattct	tgctcacttc	tataacaaac	ttacataaaa	tggattttaa	840
aattcctact	cacagcctaa	aacttctgga	gttactacc	tttttttcaa	attcatagta	900
agatcacctg	tgtattttat	attttagtaa	agccaattat	gaagtacaag	tatcatacac	960
gtacttttga	gctactatta	tttgaaaaaa	atctgccaaa	tagcatcttt	aggatatatt	1020
tacattttca	ctcatctaaa	aagtatacaa	aaataaaaag	tggaaaaagg	tatcttctga	1080
atgttcaaga	gcattctata	gtgccaaata	ataaagcacc	atttttttct	tcataaccag	1140
gattaaaaat	catatatact	gcagggcaga	catacatatg	atagcttggtg	ctgattaatt	1200

16

taacccatt	tgtaaacaga	tgaaaatttt	attttcttat	ttcatttata	agatgggtca	1260
atgtattggg	aggcttcttt	tttattacag	aaagtgtata	ttggtatata	ataaatgaac	1320
ttttcaaatg	aaaaaaaaaa	aaaaaaaaaa				1350

<210> 29  
 <211> 1766  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (1743)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1748)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1749)  
 <223> n equals a,t,g, or c

<400> 29						
gctcagcgcg	ctgcccggct	ggggacccgc	gcacctgcag	cgcccgcctgc	tcggccctgc	60
atcctgcctg	ggcatcctgc	gcccggccat	gacggcgcac	tcattcgccc	tcccgggtcat	120
catcttcacc	acgtttctggg	gcctcgtcgg	catcgcgggc	cctgggtcgt	gccgaagggg	180
cccaaccgcg	gagtgatcat	caccatgctg	gtcgccaccg	ccgtctgctg	ttacctcttc	240
tggctcatcg	ccatcctggc	gcagctgaac	ccctgttctg	ggccccagct	gaagaatgag	300
accatctggt	acgtgcgctt	cctgtgggag	tgaccgcgcc	cccccgacc	aggtgcccg	360
ctctcggaat	gactgtggct	ccactgtccc	tgacaacccc	ttcgtccgga	ccctcccca	420
cacaactatg	tctggtcacc	agctccctcc	tgctggcacc	cagagaccgc	gaccgcagg	480
cctgcctggt	tcttggaagt	cttcccagtc	ttcccagcca	gcccgggccc	tggggagccc	540
tgggcacagc	agcgcccgag	gggatgtcct	gctccaatac	ccgcaactgc	ctggagtttg	600
ccctctttcc	caaggagatg	ctgctgggga	gctgggtatg	gtggggctct	tccctttaca	660
gacggggcag	atgccaggac	tcagcccac	ctgaggagga	cacgtgtcct	catggagagg	720
gtgctccggc	ccaggcgggg	gagtcgggtg	ccagtcagca	gctctgccac	catcctgctg	780
ggaaactggg	gggcctctat	tgggttatag	gcaaggcctt	ttctctggca	tgggaattgt	840
aattttctga	cacgtctaga	tgtgaaat	ctgaaaatgt	tgaagcagag	aaacattcac	900
acacaaaaag	caacatagtc	atgtgggtcc	agatggcctc	agtcctagat	gttggcacc	960
tttgctgtgt	ctcctcagag	tatcctgttc	cgctcctgc	cacctggacc	tcctcagtg	1020
gatgtcttcc	ctcccccgac	cccagcctgt	cagtcgcagc	acagtgcagg	tttggctctg	1080
acttgggcct	ttggctgcag	tgggggtgga	tttcagagcc	tctcatggca	gcatctaagt	1140
gaccagagct	gggatgagag	aggggaaggg	gcaatgtgag	tggcgctatg	ggacgggcca	1200
gccctgctcc	tgagccagcc	ccgccctctg	ccccctggcc	ctgggctctg	tgctagggat	1260
ggtgaagaat	gggggctg	carcctgcag	gagtgggaag	caacacgcag	gggtcccgga	1320
cctctcmagc	cttgccctya	cgcttatccg	agctcccagt	gtgggttagca	cagagctcac	1380
ccaccttgcc	tggctcccag	ctggggcctg	tcctcactgg	tgctccaggg	gaagaaacga	1440
cagcctcact	tctgtatgga	ctgctgatgt	ggcctgccat	cctgttcagc	gggcattgtc	1500
tttgaggcag	caggagacta	ggatgcctct	cactcacatg	ccagttcctg	gctggccagc	1560
tgctcagggc	tcaggctggg	gcctcccatt	gacatcctcc	ccctacactc	cctctctgag	1620
cctccgctcg	ccctcctgtt	gggtaagggt	gttgagtgtg	acttgtgtgt	aaaacctggt	1680
tcatatataa	taaataatgg	tgatgaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1740
aanaaaanna	aaaaaaaaaa	aaaaaa				1766

&lt;210&gt; 30

&lt;211&gt; 2790

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 30

```

ccgtcttggg ctactggcag ctctctctct tgggctcctg gctgccaggc gttgggtgcca      60
cttcttaaaag gcctggaacc agggaggaga ggaaatgcta ttgttgtggg ctttctccgg      120
gggtctgtgct gtgcctgcta gagcaacccc tgtaccagc tccttttgtc ccaggggccc      180
ctccctctgc cccaagcagc cagccagtct tgcctaggcc aaatgcacaa gctcagaata      240
gatctgatgg tgagctggga agctgtactc agagcagagc aaatgagggg gggggcgctc      300
aggacccagg ccctccatgg gctaattgta gtggcagcca tgcctcatgc cacaccttct      360
tcgcaaaactg atggaccggg tgggcctggc ctgagctggg gccacaaatc aaatcaaggg      420
ctccagcatc cagcctgtgt gttctgtaat caggtgagc ccctccctg aaaacgaaag      480
ggccccgggg ctggcaatca gggaagctc caggtgagc ggctgtggca caaacttctg      540
gaaggctggc tgactggaat gcagggaaaa cggcagtagc tgggaaagga cccacccatc      600
ttctgtctgc tgtaactgct gagccactcg cagtcccagg atccgctgcc accacgtctg      660
ccaggcccat ctgaggtgcc actccctgag ctttggggac agttggcaga gaaggcctct      720
tgtgtctcag ctcccccgca gtccccagcc cttctgcctt tctccccga cactgctgca      780
ccagagtgaag agggctatgg caagggggtg tcatctgagg agtattaaga atgcagattc      840
ctgggcctgt cccccaaggt tttggagtca gtaggtccaa gggccatact tttgagaggg      900
gtttgggtta agtatgaggt gaaatgggag atggtcagtg tggagagggg tgcacccact      960
caccagggtc cgcaccagct gctctgcccc ttgggcatcc acccagtgct gccatgccac      1020
tgccaggcac ctggcctgct gggaaccccg cagccctgta agcagtgctc cgaggcacccg      1080
gcgctgcagg tacttctctc tgatggccaa gagcatcgtg acccttcagg gccagaagga      1140
gggcagagcc atgggcctgg gcctgctttt ccaggatcct gcaggaacga gcactggcca      1200
gagagggccc agctgtagcc atggctcagg caagcccctc agcccttgcc cccatccctc      1260
ggaccaccca aactgcacac acagctcctc ttaccgtagc ctccgtttat gggccttgct      1320
ttgggctttg caggctctgg gctcagggct ggagtgcgct cttgggtccct ggtccctcgt      1380
ccacaggggc aggcctggga cccagctact ctgtccaggc cactgtggcc agagctggaa      1440
ggcagggcag agggaatgtt ccctgcaccc tggaaagggg agttgagtca caagaggtta      1500
aggtgggtcc aggaaggcag ctgctcttag tggcgcccta ggagttgagt acagtgagga      1560
gggtggagga aggtgctgag cttagccttg tgccctgccc ccatctcccc aggcctccag      1620
cctctcccg ctgcctgcca cccaaagaga aatcacaggg gcggggcagg aatgcaaagt      1680
gttttctcag aacagctgaa acattccgaa gagggaatgg atggggagaa tgggtcaatac      1740
acataagacc gtgtcccaag gagctgattt ccaggcccct gaggactgga gaccgcttca      1800
cccctgcact tcagacaccg tttgtcccc ggggcaaggt ctcttactc tgagcccagg      1860
ccgttccccct tggcttctct cgtccacca ggctgcactg cagtgatggc gcgggaggca      1920
ccagctctgt ggctgtgtc cagcagctgc gggcttgaag gaatagccag agaggagcac      1980
ctgaacccca tgggcttgga cttcttgagg ccccgctggg atttcttcgc tgctctagct      2040
ggcaggacac atcccgccct cttcttccac ccattcccc atgtggctga agacattcca      2100
acaatggggt gggcccataa tagttagccc tcagtcagtt cccggagcac agccctggga      2160
gggggctatt tctctcccca ctgaaaacat ttcaaagctg agttacttgt ctgaggcctc      2220
atccctcgga agcctgtgta ctccagagtc tgagcccccg gctagtacct tatagagagg      2280
gggctctcca aaggggctgc tggggcatgt gtgctgtgg cagaaaagag gagaccctgg      2340
aattcagcac cctgggtgcc attcccagcg tttagtttct agaggcctca gtttctccat      2400
cagcttatgg gatccttgtc tttactgaca agaatggaat agaaatgtaa aagtactctg      2460
aaaagcaatt gccctgtaac ttatctagaa agaaaagacc ctgagactcc agaactgtgt      2520
gttgccatag ccccatatgt gtgaattctg caactagcca aggctagttc ctttcaattc      2580
catttaaaaa acaaaaacca gcaggtgtgg tggctcatgg cgtaatgggc ctgcccaatg      2640
ctttgggagg ccaaggcagg tagatcgctt gagcccagga gtttgagaca agccctggca      2700
acatagtgag atccatttc tacaaaaaaa aaaaaaaaag gaattcgata tcaagcttat      2760
cgataccgtc gacctcgagg gggggcccg      2790

```

&lt;210&gt; 31

&lt;211&gt; 1417

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



18

&lt;400&gt; 31

tttttttttg	attaaaaaaa	tttaaaaaat	tataaaatga	tgtcctatat	gagttttaata	60
catgacgttg	gaggagcata	gagatagacc	tagactaggg	atgtgtatgt	gtgtgtgtgc	120
atgtgtgtat	gcatgcatgc	ttatgcatgt	gtgtgtgcat	gcatgcttgt	gtgtgtgtgt	180
gtgtgtgtgt	gtagagcctt	ggatcatccc	acagagcaaa	gacacaggag	gggtggcacat	240
ggaagaacaa	gtgactccac	cctcccttgc	acagttaaaa	tctggccaag	tgagaggggga	300
gatgggagag	gggagagggg	agaaaggaga	agaggcactg	actggagggg	ctgaagcttt	360
gtccctcctg	ggcaggcggt	ctccatccac	acccctcttc	ttggatagag	aggataagca	420
ggccaaagat	gcacgaaacc	tgagtccac	tgtagctcca	gacttctaga	aaagtcaaca	480
gccctgtat	ctctagctga	tcctctgttg	ttcaatgtct	gcattaccgc	actgggagac	540
acttgacaga	ttgggcctgc	cgcaggccat	agcagacatt	gggcagccct	agaacgaagc	600
tgactgtcct	tggaatgtgc	cacaggggtg	tgacgccccg	gccaactcca	gtgctgccta	660
aaatggcctc	ttgcaacatt	cccctctctt	catctttaat	cagggaactg	aagccacaaa	720
atggcaaata	cacagtctct	gcagtcgttt	tgagtattgg	agaaatcgct	ctggccatct	780
gttttgtctc	cagcatgttt	ctcacggaat	atccacggat	atatccatgg	atataacaga	840
catctgcca	aggcagagct	tggctcttga	gaactcggca	agctcagtgc	ttgcctggat	900
tcctgcctca	tgtcccatcc	agtgtttgga	gaaaagctct	gagagaaaga	tgaatgtctg	960
aggccacaca	gcctagaagt	agtcaagagc	acaggctcta	gaactagccc	cacgtgggct	1020
gaaatcccag	caccagcgcc	tgccggctgt	gtgatgtagg	agagcttctt	accagctctg	1080
tgctcactt	gtctcacttg	taaaatgaga	ataagaattg	gccgggctcg	gtggctcacg	1140
cttgtaattc	cagcacttcg	ggaggctgag	gtgggcggat	cacttgaagt	caggagttca	1200
agaccagtct	ggccaacgtg	gtggaaaccc	cgtctctgcc	aaaaatacaa	aaatttagcca	1260
ggcgtggtgg	cgggcacctg	cagtctcagc	tactcaaaag	gctgaagcag	gagaatcgct	1320
tgaacctggg	aggtggaggc	tgctcagtga	ccaagatcac	accactgcac	tgacgcctgg	1380
gtgacagagc	aagactctgt	ctcaaaaaaa	aaaaagg			1417

&lt;210&gt; 32

&lt;211&gt; 1906

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (617)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (940)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1461)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1901)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 32

tctggagctc	caccgcgggtg	gcggccgctc	tagaactagt	ggatcccccg	ggctgcagga	60
attcggcacg	agcccagctt	ctgggtggtg	ctggcaatct	ttgggtgtcc	ttggcttgta	120
gatgcatcat	cccaatctct	gccttcattt	tcattgctga	ttctccctgt	gtgtgcatgg	180
ctgtctctgt	gtccaatttt	tcctttttta	taaggacacc	artcatattg	gattagagcc	240
cacctaatg	acctcatctt	aacttgataa	tctgaaagca	ccctatttcc	aaataggggtg	300
acattcarag	gggtgttagga	cttcaatatc	ttttgagggg	acacagttca	accataaca	360

19

cctaccaacg	gtttctggaa	tcattctctca	aaataaaacta	tctaaactct	aatctttgtw	420
tcaggggtcag	cttctagcag	aactcaatgt	aagacaccct	tttaaagatg	gtacctcaga	480
ggtacagaca	gattgtgaac	cttccctagc	acagtgtaac	aagtccttg	caaaaatcct	540
gattttgagtc	aacattgtaa	tttcttgctt	aaatctaaga	atatgcctyc	cagcttcttc	600
caagactatc	tgggggnagt	tgkttctagg	gtgggtcaat	ttattcgctt	tgaacgctga	660
tggccgtgca	ccacaccgtg	gctcaatggg	tgttgaggac	acgttcacat	gcaataaatg	720
cacagggcct	tgccattagg	tggcattagg	ggaaggacac	ttycagcata	ctgcagaggc	780
attctcttcc	agcctgggtg	cctgtcaagc	acctgctgag	atgactggcc	cagacggaga	840
ttgtgggggt	gctgagccag	ccctgtcccc	tccttatctc	cgggaagatg	aataaatgcg	900
tgtcagaagg	gaggggtgct	cccattctcg	grgcaccggn	gttgcccttc	agaaaaacat	960
tgccctgcaca	ttttgtagtc	ttgaaatgaa	tctgagtggc	aattcaagcg	ggcagagctt	1020
gtttttggatt	ttagacagtt	ctacctgcgt	gctcctcttc	tctgctccag	ctctgacatc	1080
tcggctccac	atacagtggg	ctgaagtggc	atacggaccc	tgagaagagg	agaggcaaag	1140
gwracagctg	tggctggccc	ctcctgccyt	ctgggtctct	tggctgggtg	gggaagaaca	1200
aacacagttg	tggtttccagg	tcacagctgc	cagggtctca	cctgtggggg	gggtggctcc	1260
agttctgttc	tgagtttaaga	aatgctgcaa	atacgttctc	cttraaagag	mcctaggaat	1320
tgccatttct	ttctgcagct	ttctgtagcc	aataagcatt	tgaggaaactg	ragaagggtt	1380
cagccctgaa	ttgcaaggga	aaactgtgtg	agtgtgtttt	agtttaagaaa	aaagttaatt	1440
ctagttagac	ctgcttggtt	ncaaaacaga	tgtataggcc	agacagatgt	acagggatga	1500
ccttgacttt	cttttgtcat	tgaggaagg	tgggttatgt	atgaccctcg	gttaagacca	1560
ataggaggcc	gggcgcagtg	gcttacgcct	gtggtcccag	cactttggga	ggccggggcg	1620
ggtggatcgc	ccgaggtcgg	gagtttgaga	tcagcctggc	caacatggag	aaaccccgct	1680
tctactgaaa	aaaaaaaaata	cagaattggc	cgggtgtggt	ggcatgcctg	tggtcccggc	1740
tgctcgggag	gctggggcag	gagaatggct	tgaacccggg	aggcggagggt	tgtggtgagc	1800
cgaggtcgcg	ccattgcact	ccagcctggg	caacaggtgc	ggaactcggg	caaaaaaaaaa	1860
aaaaaaaaaa	raactcgggg	ggggmaccca	acccgggtcg	nacatt		1906

&lt;210&gt; 33

&lt;211&gt; 543

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (367)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (376)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 33

ggcacgagaa	aatattgact	cctatctggc	cttcatcaac	tgacctcgaa	aagcctcatg	60
agatgctttt	tcttaatgtg	attttgttca	gcctcactgt	ttttacctta	atttcaactg	120
cccacacact	tgaccgtgca	gtcaggagtg	actggtctct	ccttgctctc	atttatgcat	180
gtttggagga	gctgattcct	gaactcatat	ttaactctca	ctgccaggga	aatgctacat	240
tattttttcta	attggaagta	taattagagt	gatgttggtg	gggtagaaaa	agagggagtc	300
acttgatgct	ttcagggttaa	tcagagctat	gggtgctaca	ggcttgctct	tctaagtgc	360
atattcntat	ctaattctca	gatcagggtt	tgaagctttt	gggggtcttt	ttagatttta	420
atccctactt	tctttatggt	acaaatatgt	acaaaagaaa	aaggtcttat	attcttttac	480
acaaatttat	aaataaattt	tgaactcctt	ctgttttaaaa	aaaaaaaaaa	aaaaaaaaaa	540
aaa						543

&lt;210&gt; 34

&lt;211&gt; 1452

&lt;212&gt; DNA

20

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (283)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (596)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (607)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1275)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1284)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 34

gaattcggca	cgagattaag	ttgtgcactt	taattgggtg	aattgtacat	gtragttata	60
tatctractg	tagttgtwat	taaaaaaca	caggaggcca	tgtgggctgc	taggagtagc	120
aatgtctgtg	cccagccagc	aggtagagac	caggggctgga	cagagkagta	tgggctgtgc	180
tgcagattat	ttgtggtacc	caactgttgc	ataaaacagg	gtgtgatctc	ttgcattgct	240
atgcatgagt	ggattcccag	taaattgtgc	caggctgcct	gantgatgtg	tggcttgtgc	300
tttggatcgt	aatgcttacc	tatgctactt	aagttacata	ccctgtggcc	tttgtggcca	360
ggactgtggg	ctactacctg	kagtgattcg	ttaggggaaa	ggaccacag	cctgtgcagg	420
aggaaaaaag	catctctgag	tacagggtgg	atgagctgga	tgagctgccg	ggcaagagcc	480
acgcacaccc	agggtggtgag	tcttaaggat	aagggtggaat	ttgccccata	gctgtcctgg	540
acagaaaactg	cccagagaag	aatgaatgga	ggacataggg	ctctgtgggc	ccaccntttt	600
ttggganacc	tgtgactggt	cctgttacca	tgtcaactta	gccccaaacc	catctctgat	660
tgacttggtt	gcttattttg	gcacattctt	gctccacaca	gccacataca	tactggctgc	720
tcctcsaagg	ccaggcagat	gcagcagctg	ttgggccagc	aaagaggaar	gtcctggaag	780
gttctggcct	gaacgctgca	tctgttgtgt	gacagccaca	actgctcagg	cttccttgtc	840
tgtgggtgca	ctgtggggag	gagtgttatg	ataagaacat	tggctctcag	tctccctggg	900
gagaagtttg	gcctcacgtg	ggatttgggc	gttgccctta	ggaaggctct	ctgcatgtct	960
agttccagtt	tgtactggga	agaattaaaa	aagtctgccca	gcttcttttag	tttgtcctgt	1020
cttttgtgat	gattctttct	gagatcccct	cctatcagct	caggagtggg	attttctgga	1080
gaaggaaagt	gtttttctgt	tcctcactgc	tcaccttggg	gcattcagga	acatgggcct	1140
gatgaatttg	cttgaaggca	gtctgtaatc	ccatcacttt	gggagccaaa	gargcggatc	1200
atttgagctc	aggagtttga	gaccagcctg	agcaacgtga	caaaaccctg	tctccaccaa	1260
acaacaacaa	caacnacaac	aacnacaaca	acaactacaa	caaactgggc	tggatggcac	1320
gctcctgtaa	tcccagctac	ttgggaggct	gagatgggag	gattgcttga	gcccaggagg	1380
tcaaggctgc	agtgagccat	gattgcacca	ctgcattcca	gcctgggtga	cagagggaga	1440
ctgtttcaaa	aa					1452

&lt;210&gt; 35

&lt;211&gt; 2908

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<220>  
 <221> SITE  
 <222> (1653)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1655)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (2850)  
 <223> n equals a,t,g, or c

<400> 35

gagctctaga	ctgatcgtag	taagttttgt	gacttaacaa	tgaaagaaag	tagaaagatg	60
ctttgggttt	tcaaaatggt	gtttttttaa	attgttcttt	gggtgaattt	actcagtgct	120
gctttgagtt	gtatacagaa	acaaatgttg	ggaattgctc	ctcagaaatg	tgttcctaag	180
ttgtgttttc	aacttttcat	catgagggtga	agcattaggg	aaagagattc	tttcgatttg	240
tttaataatc	taattataca	gaccaaagtg	ctttacgttt	ctgctctatg	ttaatgtttt	300
agaatgggtga	ttttgctgat	taatttagac	ctggcaattg	aaataatgtg	ctcagaataa	360
taacatgggt	atagttcttg	tatgataaag	tattcaattt	cagaatagtg	ttggaatcct	420
gagtttgaat	aatgtgttgt	atttagaaac	atagccctat	ctgtttttaa	caaataattt	480
gttggccgat	tgtccatggt	tgagcatgac	aaaaatacct	cgtcgaaagg	caagcttagg	540
taactgctgg	caaaacactg	ggtgcactat	ttttctggat	aaaattttata	gttattttct	600
atattaccct	tcaaaaggga	tctcttcagg	ttaaaaatca	cgcttatgct	gaagtcctta	660
tctgggtgta	actaaaaatc	tcatatgggt	tcataaccga	ggtcactaat	aattcatttt	720
tatcatctgt	aaaaatttgc	tcaaaattcc	aaaaaaatat	tgatttggtt	tttagtgatt	780
ttgcaggctg	accccaacct	aagttttgat	aacatctggt	aagtcagtat	agttctgtga	840
cttcatgttt	taactaagaa	ggaaaattca	tagatattcc	tattagattt	tataaacctt	900
caaaagtctg	aaacttaatt	ttgagtctaa	attttctgac	actggccctt	tttaatatgt	960
taagtttttg	ttcactttct	taagtaaaaa	aaacatttaa	ttactagtta	gcccttaact	1020
gggaaactca	ggtaatgaac	tgctgacttt	tctaaagttc	tttaactgat	caattctgta	1080
tagagggata	tttatctaac	cactttccgt	attttacaag	tgctctttct	aaaaagggaat	1140
aactattata	gctctatttc	cccaatctct	ataggactca	tgagagattg	cttgtgtaaa	1200
tataaaagca	ccatatgtgt	tcttaactcc	tatggctgct	tgaagctcat	gatgaaaaag	1260
tctttttgtc	agttttaatt	gttaagtaca	gaacaaacaa	ttgtttgggt	atggcctggg	1320
tgaaagagag	catataaata	tatcccagtg	gaactcacca	aagaagacca	caactcagaa	1380
attattgcat	tttctcatta	tgtgttggtt	ttgatttgct	tttgttttta	atgcagctct	1440
tttaataata	agattcttga	tacagtgaag	tctcttattt	caagtgtgaa	ttattcttca	1500
cccacccctt	cccctgccat	tgtattttcc	atctgtttca	aggagtttca	acaattttaca	1560
ttgcatcgta	tgcagtaggt	actgcttttt	cagaaagacc	tggaaaacat	acctgctatg	1620
aatattttgt	tcagatgtag	ccattttacct	ggntntcaag	gttgccttct	gtggagagga	1680
tcttagacaa	aaatcttcct	tgtattttact	tgggttaagt	gaagtccaaa	ttcttacagt	1740
atgctatttc	aggattttctg	atatttaaaaa	agaaaaaaca	aaatctttat	atctcttatt	1800
aacactttcc	ccaagaaggg	ttgtgctggt	atttattttc	tattataaga	aaagttcatt	1860
ctttaagtag	tttctttttac	ctctaattcta	atttcatacc	aaatacctga	tcaatagaaa	1920
tgatatattt	aagcagcaaa	gatttcctaat	ccatcattat	gaaaagtgtc	agcatactta	1980
gtagtgaaca	gataaagtca	atttgaatat	aattccactt	tgtttttaga	gactaaatta	2040
agattcaatt	aacattatcc	tatgaattct	gaatgtgata	atgtgattca	aacagtcaaa	2100
ttttattaag	ctcttagtaa	ctcaggatag	cattccatac	taacctcaag	ttagcaaaac	2160
aaattagtta	aacagcttgg	ttcttagcag	actgcttaaa	agatcaagaa	aattttctca	2220
tcttttcttt	ctacttagaa	acattgcaag	aaaccttgga	cagtcttcac	cagacctgcc	2280
atgattttat	aagatttagg	cctcagtgac	atgctcctga	aagtttctctg	ccagccatcc	2340
aaactaagca	tccactcatt	ccatcttccc	aaagtcactc	accgataaag	gtagcatcct	2400
taagtttcatt	tttgaaagg	ggaggaggat	ctccccctgc	ccaaaggaat	ttttttatca	2460
gaataccttg	aaaggggggt	atataaaattt	ggaaaactta	atttcttggc	tgtgtttgat	2520

22

aacagttcct	atgcatgggt	tttaatgtga	ggtaaatttg	tttctttctt	cagaataacct	2580
ctctccaccc	cccaccttat	tcttcctctt	taatgaatat	ttttattgga	gctcaaacctc	2640
catgacttac	gtgctcacta	agttttcttt	tttccccttg	tttactctgt	ctgtatgtat	2700
gtcaaaagct	ggcaaaacct	ctaaaactgt	caagaaaatg	cttgaaagtt	gatttgtcat	2760
agtgcgaatt	catgataaaa	tgtcttaatg	ttatttggat	atgtagtaca	tagaaacaga	2820
aaaataaagt	cattttttata	acttaaaatn	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				2908

&lt;210&gt; 36

&lt;211&gt; 953

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 36

aattcggcac	gagaaccaag	gtactttcag	ctgcagactg	accatacccc	tgccacactg	60
gcgttgctgg	atgagatcag	cacctgccac	cagctcctgc	acccccaggt	cctgcagctg	120
cttgtaagc	tttttgagac	tgagcactcc	cagctggacg	tgatggagca	gcttgagttg	180
aagaagacac	tgctggacag	gatggttcac	ctgctgagtc	gagggttatgt	acttcctggt	240
gtcagttaca	tccgaaagtg	tctggagaag	ctggacactg	acatttcact	cattcgtctat	300
tttgtcactg	aggtgctgga	cgtcattgct	cctccttata	cctctgactt	cgtgcaactt	360
tctctcccca	tcctggagaa	tgacagcatc	gcaggtagca	tcaaacggga	aggcgagcat	420
gacctgtgta	cggagtttat	agctcactgc	aaatctaact	tcatcatggt	gaactaattt	480
agagcatcct	ccagagctga	agcagaacat	tccagaaccc	gttggtgaaa	aaccttttca	540
agaagctggt	ttaagaggct	cgggcagcgt	cttgaaaatg	ggcaccgctg	ggaggaggtg	600
gatgacttct	ttacaaagga	aaatggtagc	agcttcagtg	agaaactgcc	cttacaacaa	660
gtcccttctc	tgctgtcaat	ccaatactgc	tcccaaattc	tgttttcagt	gttcatttcc	720
ctcaaggcag	gcgctgggct	cccacgaccc	ctcaggacag	atctggccgt	cagccgcggg	780
ccgctgggaa	ctccactcgg	ggaactcctt	tccaaagctg	acctcagttt	tcttacaaga	840
acccagttag	ctgatgtttt	attgtaattg	tcttaatttg	ctaagaacaa	ggtaataagt	900
aaatttttaa	aaagcctttc	tgctgggttg	gattaaaaaa	aaaaaaaaaa	aaa	953

&lt;210&gt; 37

&lt;211&gt; 3864

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 37

accacgcgct	ccgtttctga	agcaatgtta	atcctactag	ccaagcatat	cacttagtcc	60
ccactgtgag	atgagggata	tgtgcttaaa	ttgtgaaaca	aatatatgag	tcagggtattt	120
ttcctttgag	tccaagtggg	ttatgacttt	ctttcctgtg	ttctttgtat	atgtgggagt	180
tttataattt	tttatcaaga	atgaaagggt	ggcctgtggt	cttactgggt	caggctgtca	240
catttctctc	tgttgcccag	tcagggtgcta	tggcatgtgc	tgcttctggc	gtagtgtact	300
ctgtggatgt	accagcatgt	tcttcaagggt	catgactgat	ttccagacc	tttggaattg	360
agataaatgt	taaatttgta	gctatctctg	aatttcttcc	agatactttt	cttcatttgt	420
ttgtttgtag	ggtaaacata	cctgatagca	gcaatttaag	cataccctta	gaatgaccat	480
gtatggccag	tgcacctgaa	tgtgtgttcc	aaggtaggga	atccaggaat	ggccaactcg	540
gagattcatt	ccttactatg	ataaatatct	gagccccctg	ctcatcctgt	ggaacatggg	600
cttattgggg	attaaggccc	tgagtttttag	gttaaatgaa	ggttaccaga	tggaggtcat	660
tagggggagg	gtgttaaatg	aaaatgcttt	ataaactgca	tgctgtttgc	aagcagttgc	720
agttttcctg	cccagcccgc	agccactggc	catgcagtca	tggtgtccag	cctgccgcca	780
ctggaccatt	tctgtacata	aggcagttct	cctgtccgcc	tgccaccagt	tctccactct	840
ctccccatat	gtaagcccct	agtaaaccct	atgtctcatt	tgctgcctct	gggtcttttc	900
ttcagcctct	tgaacctagt	gccttccctg	ctgagggttaa	taggggtaca	gcacaacagt	960
gttgtaacac	agaaagtgat	atttacaggg	atatctctct	cacaatatct	cttaggaaag	1020
gtaaataaaa	tgttcacaaac	ttgtaggtga	gtaattcctt	agataagttg	tttcttaact	1080
tgggaggagt	ttgggaagga	acctaagcag	gctgcagagg	ctgggcatgg	gagcttgtca	1140
tggctggaag	ttgaaatggg	caactccagg	cagatctcct	ggggcaaacg	agcctccacc	1200

23

accagtagcc	cttcctttct	gtctttcata	ccccactgct	ccatctgaag	cctgaacccc	1260
ttccagaaaa	ttgatggata	gatttttttt	ttcggctata	tatagtttta	gaggttagaa	1320
ctagatataa	ttcagtctag	aagatttctc	cttccccaga	aatgattggt	tttgtcaaaa	1380
gccccgcaa	aatagtagcg	agacttagac	tgagttcact	catcactaac	aatttaacttt	1440
ataaacattc	aacaagtagg	acaactatta	ttactgttac	tcagaaccct	tcgctctgta	1500
tatacagttt	gattttaagat	gccacattta	catggcattt	tcaaccttca	aactctagca	1560
gattttaaaa	ctaggtggat	gaaaaatagaa	tcatttcta	aaatgtagtg	tgtagatttt	1620
gaaaaatcat	ttggtgagca	ggatctctgt	aaagttatat	gggccacgta	tacaagacgt	1680
aactgaagaa	aattaattca	acagagcatg	ccgtacttga	acgacataga	gattttactcg	1740
aactgaacta	actcaagctg	cagaactccg	agcaagcctg	gatttgtaaaa	gtctgggtga	1800
aaatagatgg	agtatgccct	gactgaacct	ctgtactgcc	ccacatgctt	atacaggtgg	1860
gggattggat	ggctgttagg	tgatcattgc	attctctttt	ggatccctat	tgagaagaaa	1920
tgataagaga	gggaaaggat	atggggcaag	aacagtctga	aaaagaaagg	ataaagttct	1980
cagactctct	tcacactcta	agaagaactt	tctgaaaagc	ttggattagg	tctggcaatg	2040
gatataataa	gcaaaggact	cttggaaatg	gttcttggct	cttagcccca	cctctgactt	2100
tgagcaaata	agctgatttc	tctgcctgta	aaataatagt	ccctctgata	ttaataactta	2160
cctcatgagg	ttatttagag	gatagtgttg	gtaataatgc	cttgtgttta	catcattcct	2220
ttcacagaga	gctcaaagca	ctttacatgc	attgagagag	aagcttctcg	tgaagagtaa	2280
atagaagtgt	tcactttttg	gaaatgaact	tagggcataa	gagcctgaat	ttaatgcatt	2340
gcaggaagaa	atatggtaaa	tagtgaacca	gtgggtcaac	tgaatttttt	gttccactaa	2400
gagtccccct	ctggctcctt	gttttggtga	ttgaggaata	tggtgagtc	ctacacctgg	2460
atgggaaatc	ccacatatgc	aattggaatg	gtctctcacg	acacatgcag	agattgaaga	2520
acagtctgga	cattttttga	taacgttctt	tgggccttgg	tagtagctga	aagacacctg	2580
agaaatctta	gctcagagct	acagaatgac	actaatggat	cccagaaata	gaaatgtaga	2640
tgtggagtgt	tttatctgtt	tatttcacct	caattcaacc	aatactcctt	gagtgccttt	2700
tatatacatg	attttgagtg	atgtggagaa	taaaagagc	cccacatgct	caggaaagt	2760
aaccctgggt	ttagcaagga	aaagaagtag	gattttccaa	tagataagt	caccgggtat	2820
gtggaagtgc	agaaaagctt	cccagtattt	cagcccatct	acttgcccat	tctcaaccat	2880
gtattactca	tgtaccaagc	agtatgctgt	tcacagagag	atccaatctc	tgcttaggg	2940
atccttgggg	aaaacatgta	caaagagata	gttttagcac	attctagtaa	aggcagtgt	3000
caagggcacc	tagccttacg	tggcaattta	gggaaggtac	cttgaggat	gagacttctc	3060
ctaaagtctt	aagaattgaa	aagaacatgg	gaagggaatt	ccaggctggg	agagttagtat	3120
gttcatacgc	cctcagtgtt	taaccttctt	tgaacaaaaa	aatggccaac	tacagaaagt	3180
ttgggtcttat	tgtagcctaa	attgtactta	ggggtacgag	tgagaaacag	ggattaagat	3240
aaaggacctg	ttttgctgtc	ttgtttactg	ttgaatagta	gtatgaagta	ggctcctgaa	3300
aactatgttt	ttggggaaaa	aaaaaaaaag	actgaatgat	atgttgggtt	taagtccttg	3360
caggcaggct	atccaggtaa	ataaacatgg	aagggtgatg	gaggtaatct	gggctggaaa	3420
tacagatttg	gaagtcacc	catatcagtg	gtgtttaaaa	tcaagagcaa	atgaaattgc	3480
acaaggagaa	tatatagaat	gaacaaatta	ccatgggtga	agccttgagt	aatacagaca	3540
tttaagaagc	aaacaaaaga	caaggaaacc	atgagggaga	ctggaaagga	aaaaacagag	3600
aaataagaaa	aatgagagg	agagaattga	tacattttcc	tcaggtgtgg	cattatggag	3660
ttcactgggt	tctcatcaga	gaagtttcag	tccagtggcc	agggcagaat	gacattgtgt	3720
cttggttttaa	agtaaaggg	tagggtaaga	aagttgagaa	gggttagcac	aaacctctct	3780
ttcaagtcac	ttgccttaga	agagaaggaa	gggtatggtt	tctggggtgc	aaccaggtt	3840
caagaagcaa	aaaaaaaaaa	aaaa				3864

<210> 38  
 <211> 1411  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (1395)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE

24

&lt;222&gt; (1397)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1401)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1408)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 38

ccggtccgga	attccccgggt	cgacccacgc	gtccggcgtg	aaccaccgtg	cctggccgga	60
agtctttaaa	aaataaagt	attctactct	tctaagctta	cagagaccag	accagggtgaa	120
tgtaactggg	gaaaatcaag	atggtacctc	tctgcattat	cccgccagac	actgtatttt	180
atgcattcat	gtctaggata	cagtgtgaaa	attaaaaagt	ttagagggca	gatgcaattg	240
tggcaagtga	cctgcccaata	aagcagggtgc	agctatagaa	gctggcatag	gtatatcctt	300
aatggtgctt	tctccctggg	cttgtctttt	tggtgttttt	ttcccctata	ttcagagctc	360
cttgagaagt	gataaacacc	tccagctttc	taacatcctc	cccacaccat	ctcaccatat	420
ccatctccca	gcatccatct	gcattcagct	aagggcgggg	aactgacctt	gtgcctgtgt	480
tgcagaccat	ttctgaggtc	tccaccatcc	aaggaggcac	agccgtcatt	actgtcctcc	540
atgccttcag	cagccccctt	cacagctaag	gtacatacca	ccccttctgc	cgcgccctcca	600
cccctggcac	caaggctctt	tgctgcttat	gtctaaaggg	atcacctata	tttaactgcc	660
tcagtgcact	aacctctttc	ttctcatgtg	ccagatgtta	agatgaagga	ggaatacmac	720
acatactcaa	gcctcagcct	gtttagttgt	tttcaactgg	gctcgctttt	ctgggacggt	780
atattattat	agactggcaa	gcctaactcc	ataggtttac	aggaagttag	gatattttta	840
taaaacaatt	gtgtcctccc	cacattttgc	tatgttaata	tttgcttcta	acaatttgca	900
gctgtttcac	tttttcccta	tttgtctcta	agttgaaggc	tttgttgagg	gggacagagc	960
acaggaacag	ccttgacagt	ctgtaattat	tgtacagata	ttttaatagc	atataaataa	1020
gtatatctct	tttatattga	aacaaaaatg	atcagacact	gccttttgtg	tggttgctgc	1080
ctgtggcatc	ctttttttaa	aagactgtta	catattaaaa	tagtgtacat	atataaatat	1140
tacctctttt	gctgtacagt	tgtgatagag	actgaagatt	ttattttttg	tgtgcttttt	1200
ataagaaaaa	aattaataca	ctaaagaatc	ttgctgatgt	gattgtaatg	tacctatgta	1260
acttatattac	ttttgaaatg	tcttctgtat	ctttaaacct	tttattaaat	aagggtttta	1320
aaattcaaaa	aaaaaaaaaa	aaaaaaaaag	gsgggccsct	ytaraggatc	caascttgcg	1380
tacgcgtgca	acganancag	ngtcgagngg	t			1411

&lt;210&gt; 39

&lt;211&gt; 1182

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (496)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1162)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 39

ctaagctggt	tcgcctgcag	gtaccgggtcc	ggaattcccc	ggctcgaccca	cgcgctccgat	60
aagcacccat	gtctttgaat	atgaatgtat	ttgtaaaata	ccacgtttca	tgtgtgaata	120
tgtgctttta	ctgtacatag	tgctattgtg	caataggtct	tatgctgttt	tcactcaatg	180

25

tgtgctaaga	tctagcccca	ttgactcttc	tagaaatgca	gtattgcttt	gacctgccat	240
gtggcactcc	acaatgtcaa	ttgcagttta	cacacattgc	ctaaagtggg	ggacacctgg	300
gtgcccttga	ccccttggca	ccggatacag	gccacgataa	acatcctttc	gtgtgttccc	360
ttctgtgctt	gtgtggcatg	tgtacccagg	atgggcctat	aggtcacaga	ggtcagtttc	420
tcttttggtt	tccagatttt	ctttagaacg	gtgactgacc	ctcctacttg	aggccgcctt	480
ttctccttat	ccttgnccag	acttgtattg	ccagactacc	taatttttgc	cagtctcatg	540
ggtagatagt	ggtscagtgc	tttamcatac	attcatctga	tcagcattaa	tttggggaat	600
tttttcaact	agcctttctg	gtttcccttc	ctgtgcattg	cccattttct	catggagttt	660
cttatctttt	ttggtttatt	ctcaggagtt	gcttgtacat	tcttgggcaa	ttgcagataa	720
ttccaagaat	gcatatttgg	gctgggtatg	gaggttccat	ggtaatccca	gcactttggg	780
aggsccaggc	agaaggatcg	ctgcagccca	ggagttcgag	actagcctgg	gcaacatagc	840
gagacctcgt	ctctacaaaa	aaaaattaaa	aagggggcct	tgggaggcca	aggcgggcag	900
atcatgaggg	caggagattg	agaccctcct	ggccaacatg	gtgaaacccc	gtctctacta	960
aaatacaaaa	aattagctgg	gcatgggtgg	gcacacctgt	agtcccagct	actctggagg	1020
ctgaggcagg	ggaatcgctt	aaacccagga	ggcggagatt	gcagttagcc	aagggtccac	1080
cactgcactc	cagcctggcg	acagagcaag	gctccactca	aamaaaaaaa	aaaaagggcg	1140
gccgtcttag	aggatccaag	cntacgtacg	cgtgcatgcy	ac		1182

&lt;210&gt; 40

&lt;211&gt; 2457

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1622)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1713)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 40

gtcgaccac	gcgtccgcc	acgcgtccga	ttataataaa	aagtattatt	aggacaggta	60
aactcatcca	tctactgtga	gctgctggct	tggagaacga	cttttgagga	gccagtttcc	120
tgaggagaaa	tgcttttata	aagcactcat	tgctttgtaa	aaggagacaa	aactgatcct	180
aaatgaccac	tccaggggtg	ctgattttgt	ttctggctca	tgtctgccta	gtaaaccacc	240
agcaagctgc	tgaaccaggc	tggaaacaac	attgttgcaa	ctgggaggga	catagagtac	300
tgtgaaagca	gcgttccaac	acatcttcac	ttttacaaag	ggataggcag	agacttccaa	360
catagagggt	cttaaacttc	gaggggttac	aatccccttg	caaataattt	tatagctatg	420
gaccctctcc	tcagatctac	agccaacatt	ttcagtgcac	cttaggaggt	catttttagtc	480
cacgcctctg	aaaaagctgc	actccaacgg	ctgaagagca	agccatacgg	ccgagaatgg	540
ggctcccttt	gccttcatga	aagctacttc	ccccaaact	aagactcagc	tgtacgtttg	600
cttagctcag	tcacattttac	attcttcttg	gtgaactgta	ccttttgagt	acctgccttc	660
attttctaga	atcagaccta	acaaggctcag	tagaagcctg	ggcagcagcg	ggcctggaaa	720
gacgaggcag	ccagcatgaa	ctgctgttct	ctccctgacc	acaaggcgct	gtcttcctcc	780
agggctcaagt	aattgttctt	gttccgcctc	acaggaatgt	ggggagggaag	gacgttaaca	840
ctataaaaat	ctgcgtccta	ccttaaactt	gtactgtgag	aaagctggaa	acttccacct	900
gtacagtggg	tctggtttgc	atgttttagtt	tcattttgtg	gaactgcttg	tccaagagtg	960
agctcaggtc	agggcagttt	gtgcctatga	gaattagctc	agctatcagg	caggttttta	1020
gacacctttt	taaaatgtgc	tcgtgttttg	tttgttttgc	ttactgtcag	tcctgggtca	1080
atcaaagggt	tgtaaagggtg	agaattttta	tgcactgcta	tatcgcaagt	gcttaaaaca	1140
gagactggcc	cagatgaggc	attcttcaaa	tgtttgttga	aatgaatgga	caaactctta	1200
ggataaatcc	taatttggtg	gcaactgtta	tttgattttta	gaaggcaaac	tgattttatt	1260
ttagagaggg	gaaggggagg	ggaggctcat	tagcctcttg	gtagaaagag	gactatttct	1320
gcaaatgaat	aggtttccac	cttaagtagt	gacagtcctt	aacttcttat	tatggagtga	1380
gtcttgaccg	ctttccaagt	tcaatagaag	ttcaagattg	cctctcagtg	attagggaaa	1440



26

ttgaagcttt	taaagctcct	ggtctcagta	attcctcaga	ataaacctct	ttaaaagggga	1500
tattgatgga	aatgtacaat	taccagtaat	tgagggtttta	tctgagggga	tggagatgat	1560
gaaatgggtc	cttcttgga	ggtgttgga	ttttggcttt	atttttcaca	aataaagtga	1620
anccatttaa	aacgattgac	aacgattata	tagtgccatg	tggaatacaa	tagatattaa	1680
tttgtggttg	gtttttctgc	ctgctttaa	tgnaatgtat	tatgtttctg	ggttcctttt	1740
ttagctgtaa	aaatacttcg	tcactaaagc	atgaaattta	atcagcagtt	gttcttcaag	1800
ttcctgaaag	ctatarragt	ttctcatgac	ttgagtgggt	ttttccctgc	ccaccagagg	1860
agaaagccct	tgtagaattc	tgcatgttta	caagtgttcc	ctacaaaaac	tgaaccatc	1920
agctcctctt	taacaagttg	gcttttttaa	agcacgtaat	tacaatttaa	tggtattctg	1980
taaagtgggtg	ctctaggcat	aatttaaatt	cttttttaatg	actatatattc	ttcaaaactt	2040
tgaagaaaa	atgtgttctt	tttctgcat	cctttgttaag	aagactgcc	acagaggaaa	2100
aaggacttta	caaattaaga	ccatcttggt	ttcattttcca	caaagatgag	aacaaatcat	2160
ggtgttagga	aaggatcctt	agaagaacac	aagaatttga	aagcccttgg	tggttatcac	2220
tactatattt	catattttcca	cagaagtgc	ttagccaagc	tctgcatttt	gagcctgctg	2280
actttcatctt	aaaaggaatg	aaaggctgaa	aatccaggct	gctgtgtctg	tagataaagg	2340
tcaaacctatg	tttgagtctt	tcactgttgt	gtccacctaa	ataaaactga	gtaagtaatg	2400
aaaaaaaaa	aaaaaaaact	cgaggctgcac	ggtatcgata	agcttgatat	cgaaatt	2457

&lt;210&gt; 41

&lt;211&gt; 1847

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1279)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 41

gactccttag	ctaagcatac	aaggtaactg	gccttgcctt	gctgttttgc	ctcacctgct	60
aacgtgcaca	tgtcttgcca	cgtgtacctg	ctgctagtgg	gacatgctga	mttctcagtg	120
ggcctcatgg	gccagaggaa	gttgcggtgt	tccataaact	cagctctcag	aagtgcgtgt	180
tcctctgcct	ggaattcgct	tatctgcttc	aatagctgaa	tgctgcttat	ccctcaagac	240
tcaattctgg	tggaacctca	gttctgacct	ttccccagtc	tctctttccc	ttctccccac	300
aggctgtctc	cccttgatat	tttgccctgct	tgctacccat	gtccttttct	cctggtgaca	360
gtgctacagc	ctttccttgg	ggaatcaacc	ttgtctgaga	gtggacagaa	attatccacc	420
cctccatcca	gggatgaacc	aatgacctgg	ccaatcacag	tcgctgtgat	tggtcctggg	480
acaggcagga	gactcaagct	aggccacaga	gcacagtgct	tgagactgaa	actactggga	540
aaacaatatg	ctctttctgc	ctgaggcgcc	tagtggaatac	aatggaagcc	ttgaggcagc	600
tgatcatctt	tgcccttgaga	gagggtggctg	cctaagaagg	aagccaattg	agaaaaagca	660
ggcagagaaa	aaaaacagga	acagatacaa	caggcaagac	tcagctgatt	tctcatgtca	720
gcaaatacaa	gagaaggaga	ctgggtcaaga	tagagatgtc	tgatgatact	gagctcccga	780
ctccagccac	tcttgaagtc	attcctaagc	ctctcggtac	acaagccaat	acatgatctt	840
gktagattaa	aaatagtttg	actttgggtg	tgccacttgc	cgtagaaaagc	atycacgatg	900
atacactcty	ccgtgtgcty	ccacacccta	atttaaccty	ctycatagca	ctgactcgga	960
ggccctatga	gactgkgaac	tctytgaggg	caggaaatgt	atcttattcg	tctccaaggc	1020
cacgagactt	ggtgtatgat	agatattcaa	tacaggtttg	tgaataaaaa	aatgaaggaa	1080
tgtttatcta	gaaattaatg	aagcttttca	tttacttttt	tacttcaggg	cctttctgcc	1140
aagaactctt	aagatgcctt	aggatcttgg	ttgcagccag	gtggctgtcc	tccccagagc	1200
ccttggtgtc	agtcctctga	agtcattggw	cctggagata	acagggaggg	cagggtcccga	1260
ctgctgagaa	agtcacggnc	cccggatcct	accattacca	tgctacctgt	tcacttgggc	1320
ttcacccaag	gccacactca	tctccgtacc	ccttcccac	agtggtgagg	ggcaggagca	1380
cctggacatc	aagaatcgag	tgcatgcctg	aacctgtcca	ttacccatgc	ttctgcagct	1440
ttgtcatgct	tttcccctcc	ctgaaacgct	cttcccttgt	ctacttaatt	attcaaagct	1500
cagaggaaaa	aaaatcacct	gtcccaaaac	gtcttccccg	aaccctgcag	ggaaccaa	1560
tcaaaagcca	tcagaaggcc	gaggcgggcg	gatcacttga	ggtaggaggt	tcgagatcgg	1620
cctggccaac	atggtgaaac	cctgtctcta	ctaaaaatac	aaaaatttagc	caggcatggg	1680
ggcagggtgcc	tgtaatccta	gcaactctgg	aggtgacgca	ggagaactgt	ttgaacccag	1740

27

gaggcagagg	ttgcagtga	ctcagattgc	accactgcac	tctagcctgg	gaaacagagc	1800
aagactaggt	taaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	actcgag		1847

&lt;210&gt; 42

&lt;211&gt; 2597

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 42

ggcacgaggt	tacacctcac	cccctacttt	cgctgactcc	agaccccaga	ggctgagggg	60
cagtagatat	gctgcctgcc	attgcctctt	tccttgtgca	tgtctgagtg	gtgccagcct	120
gatcagattc	tgttgcaagt	ccccgtgctt	gccacaatgt	cggtcgcgtt	tcttatccaa	180
cgctgtttct	gcttctggtg	gtttgtgcta	aatgcatttt	caatcccaag	tggtagagag	240
aagaaaagga	ttgtctttaa	aaagtggctt	tgaatgtgtc	gaaagcaagg	ttgcctgctg	300
tgtgtttttg	acaattgcct	ccccagagga	gcctttgaaa	acctcacttt	tcaagacgct	360
tcaactccta	ttagaaaagt	acttaaataa	aattgctctt	ctgaccatcg	cggctctgga	420
tgccaggcta	gtgattat	gtctaactct	agtctccagg	gctagaactt	cacaccgtgt	480
cttttcacta	ccctcccaga	ggaggtgaga	ggaggtgaat	agacacttgt	cagattcttc	540
tgcctgggtg	cactgtgggg	ggatgcacac	acactggagt	ataggcttgc	taaagagaga	600
ctgcattgct	atttcaacta	gacacgttta	cttttccgcg	aaatgattac	tgagcaccta	660
tgggtgtgcc	ggtaccatgt	tagggcaaga	gatctacctc	atccttgcat	tgtaaacctg	720
tagcatccgt	ggcaatgttt	tacaccactg	ccctcccagg	gttcctgcaa	gctgactcct	780
tttattttcc	ttgtgcatgc	tgatgtgttc	aaggatattt	agtgggctac	tgtgagacta	840
tctccccatc	aggatctctc	aaccttgcca	ctaggaacat	gttgggaaaa	gtagggtctt	900
gttaagggag	gaagaggggc	tttcctgtgc	gtgataggat	gcttggcagc	atccctgtcc	960
agtagcacca	catctacctc	agttgtgaca	acaaaaaatg	tctgcagaca	ttgtcaaagt	1020
tctcctccct	attttacatt	taccaaatgt	ctctttatgt	cctttacagt	taaggaaact	1080
taggaccagg	aagaaacttg	ctcaagtcaa	acagtgaagt	ggcggtagag	ctttagctac	1140
aatttaagtc	ttgtcttccc	aaccagtggt	ccatcttatg	gatccattca	gtgggttgct	1200
cctaaatgct	taataaccaa	ttatctggga	taaacgaaag	caaacaacac	ttgctgggag	1260
cgtgggacaa	tttccatggt	gtaaataact	ccaccatggc	tggtttcaag	ctacttccat	1320
ggcgtcactg	aatgtgaagt	tgggaaggga	tctgtacagt	tggatctggg	gagctgggtg	1380
cagccagctc	agaatgcccc	caagtccatg	tgtgtgctg	atcacatatt	agtcttgctc	1440
tcctctggac	cttgttctgc	cacccccag	acacatgcc	ttatcactgt	gatgtgctgt	1500
ggggaaagca	gtcacctggt	aagagttttt	gataacttct	ctggataaac	ataaagcagt	1560
tagagtggag	ttatccccct	cttctgttaa	atgaaagt	cttttcacag	tctttccatt	1620
gaaggaatta	gaaaagtcat	gggatacctc	tggctcttaga	gcagaaatga	agttccccag	1680
aaggggggtc	tggacagaat	ggtctcttgg	gcagggacta	gatgctgtta	cttaccatca	1740
gatgtgtcta	ctgtaagatg	cttgggaggc	atcaaagtgg	tgacagtggg	gggctgttgg	1800
attttaccat	tctggaatat	aagagaaggg	gtgtggctgg	tgttttagtc	cagatagcag	1860
ctgtgaccac	tgggagcctc	attttccaca	cctgcaaaag	gagaaatagt	gatccccctt	1920
ctagtgttgg	tgtaaagagt	agatgaacct	ctgcctgagt	ctcagatgcc	tgctgcccac	1980
ttgggtgctt	agtggatcgg	ctggctgtca	acagtgtgag	cttatcaagc	ctaaataact	2040
agttggtctt	gactttttca	cctcgaccat	gactcagtgt	catgtgctcc	tagtcattct	2100
gtctgtggtc	caacgttagc	ctgggaagca	cctgggactg	agggagaaga	cctcagctag	2160
ttattcagcg	atccaggttc	tcctcctgcc	tttgatatca	cctcattata	taaccttggg	2220
aaacactttg	gtgtgactgg	aatgtggata	ctcccagggg	aagggttagga	gcatggtagg	2280
gcatttggac	tttatcatga	aggtggtagg	aaatacttga	agggttttta	gcagggtaga	2340
cacatcatca	aatgtgtgtt	ttgaaaactt	tttctcgcag	aggacagggc	cacagcccaa	2400
gcaggatcaa	accagttagg	agattgatgt	aacagtcccc	tcagaaagtg	atgaggtagc	2460
tgggtgcggt	ggctcatgcc	tgtaatccca	gtcacctggg	gaggctgaga	ctggcagata	2520
gtttgagacc	agcccttggc	aacatggtga	aaccccttct	ctacagaaaa	aaaaaaaaaa	2580
aaaaaaaaaa	aaaaaaa					2597

&lt;210&gt; 43

&lt;211&gt; 3116

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 43

ggtgataatg	aaagtgggtg	tggtgatggt	ggtaatactg	gtgggtggtg	cattgggtggt	60
ggtgggtgatg	gtgggtgatac	tggtgatggt	ggtgatggtg	gtggcgctgg	tgaccctgac	120
atgggggtcca	gtagcagtga	cagtggatgc	aggctcctgg	tgactgagga	gcattctcagg	180
ctgrggaggc	acctctgatac	cccgccactg	ctccttacc	cctacagtct	ctcagcaaac	240
ctgctggggcg	acagcggact	cagatgcctt	ctggaatgtc	tgccgcagtg	cccatctccg	300
gtttgcttga	tctgagtcac	aacagcattt	ctcaggaaag	tgccctgtac	ctgctggaga	360
cactgccctc	ctgcccacgt	gtccgggagg	cctcagtga	cctgggctct	gagcagagct	420
tccggattca	cttctccaga	gaggaccagg	ctgggaagac	actcaggcta	agtgagtga	480
gcttccggcc	agagcacgtg	tccaggctgg	ccaccggctt	gagcaagtcc	ctgcagtga	540
cggagctcac	gctgaccacg	tgctgcctgg	gccagaagca	gctggccatc	ctcctgagct	600
tggtggggcg	acccgcaggg	ctgttcagcc	tcagggtgca	ggagccgtgg	gcggaacagag	660
ccagggttct	ctccctgtta	gaagtctgcg	cccaggcctc	aggcagtgtc	actgaaatca	720
gcatctccga	gacccagcag	cagctctgtg	tccagctgga	atttcctcgc	caggaagaga	780
atccagaagc	tgtggcactc	aggttggctc	actgtgacct	tggagcccac	cacagccttc	840
ttgycgggca	gctgatggag	acatgtgcca	ggctgcrca	gctcagcttg	tctcaggtta	900
acctctgtga	ggacgatgat	gccagttccc	tgctgctgca	gagcctcctg	ctgtccctct	960
ctgagctgaa	gacatttcgg	ctgacctcca	gctgtgtgag	caccgagggc	ctcgcccacc	1020
tggcatctgg	tctggggccac	tgccaccact	gggaggagct	ggacttgtct	aacaatcaat	1080
ttgatgagga	gggcaccaag	gcgctgatga	ggaaacttga	ggggaaatgg	atgctaaaga	1140
ggctggacct	cagtcacctt	ctgctgaaca	gctccacctt	ggccttgctt	actcacagac	1200
taagccagat	gacctgcctg	cagagcctca	gactgaacag	gaacagtatc	ggtgatgtcg	1260
gttgctgcca	cctttctgag	gctctcaggg	ctgccaccag	cctagaggag	ctggacttga	1320
gccacaacca	gattggagac	gctgggtgtc	agcacttagc	taccatcctg	cctgggctgc	1380
cagagctcag	gaagatagac	ctctcaggga	atagcatcag	ctcagccggg	ggagtgcagt	1440
tggcagagtc	tctcgttctt	tgcaggcgcc	tggaggagtt	gatgcttggc	tgcaatgccc	1500
tgggggattc	cacagccctg	gggctggctc	aggagctgcc	ccagcacctg	agggtcctac	1560
acctaccatt	cagccatctg	ggcccagggt	gggcccctgag	cctggccagg	ccctggatgg	1620
atccccccat	ttggaagaga	tcagcttggc	ctggctggag	gggtcctgcg	gggtcctgcg	1680
tttctgtatg	gagctcccg	tgctcagaca	gatagacctg	gtttcctgta	agattgacaa	1740
ccagactgcc	aagctcctca	cctccagctt	cacgagctgc	cctgccctgg	aagtaatctt	1800
gctgtcctgg	aatctcctcg	gggatgaggc	agctgccgag	ctggcccagg	tgctgccgca	1860
gatgggcccg	ctgaagagag	tggacctgga	gaagaatcag	atcacagctt	tgggggctcg	1920
gctcctggct	gaaggactgg	cccaggggct	tagcatccaa	gtcatccgcc	tctggaataa	1980
ccccattccc	tgcgacatgg	cccagcacct	gaagagccag	gagcccaggc	tggactttgc	2040
cttctttgac	aaccagcccc	aggccccctg	gggtacttga	tggccccctc	aagacctttg	2100
gaatccagcc	aagtgatgca	cccaaatgat	ccaccttctg	cccactggga	taattgactc	2160
aggaaagaag	agcctcggca	gggcgctctg	cactccacc	aggaggaagg	atacgtgtgt	2220
cctgctgcag	tcctcaggga	gaactttttt	gggaaccagg	agctgggtct	ggacaaagga	2280
gtaccctgca	ttacgtggga	tatgtgtgat	caattgggga	catgcgacac	acaatgaggg	2340
tgtcatgaca	atgcatgaca	cgtacgggta	tatgtggcag	tgtgaccctt	tgacatgtgg	2400
cgttacatga	aagtcagtgt	ggcacgtgtt	ctgtggcatg	ggtgctggca	tcccaagtag	2460
caggatacat	gattgtttgt	ctatatatga	cacatgacaa	atgtccatgt	cacaggactc	2520
atggctggcc	agatgacctc	aggctggccc	aagatcta	ttattaattt	ttaaagcaaa	2580
tacatattta	tagattgtgt	gtatggagca	gctaagtcag	gaaaagtctt	ccgcccgagc	2640
tgggagggga	gagtgctcat	gcactgacca	gtccaggggc	tcaagggccca	gggctctgga	2700
acaagccagg	gactcagcca	ttaagtcccc	tcctgcctca	atcctcagcc	taccatctca	2760
taaacttgat	gactcctccc	ttacttacat	actagcttcc	aaggacaggt	ggaggtaggg	2820
ccagcctggc	gggagtggag	aagcccagtc	tgtcctatgt	aagggacaaa	gccaggtcta	2880
atggtactgg	gtagggggca	ctgccaagac	aataagctag	gctactgggt	ccagctacta	2940
ctttgggtgg	attcaggtga	gtctccatgc	acttcacatg	ttaccagtg	ttcttgttac	3000
ttccaaggag	aaccaagaat	ggctctgtca	cactcgaagc	caggcttgat	caataaacac	3060
aatggtattc	caaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	ctcgag	3116

&lt;210&gt; 44

&lt;211&gt; 3460

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 44

acgcgtccga	tgaacttgac	cggtctaaatg	ccccactttc	tcagatgggt	tctaacgact	60
ttcaggatta	gggccagctg	tgggtctact	ccttggttga	gcccattctca	cctgggatgc	120
ctgcagccag	ccctccctcg	tgatttgtct	caccttgagt	aggagacatg	cttctccctt	180
aaccttttcc	tttctgccat	aattaacata	tgtccttttc	agtaagtcca	tgctctgggc	240
aggggatgaa	gaagtactca	ctggtaatta	gctaccatct	ttgcagcagc	cctggtaact	300
tgaaaaattt	gggtctggtg	ctgttcatga	gtcctttgtg	aactgcaaaa	gcaggaaagg	360
aagtcaagac	tcctgttgcc	tcgtgcttag	caaagcagtc	cttatccttt	atactctggt	420
cttgggtttt	gtttttgtct	tgtttttatc	caggcaaatt	gcttagtagc	aaagggacca	480
aactgaaaag	gtgacaatct	ctaacttcta	aaagcagaca	ccaatcggat	gctcattaga	540
gggttaatgaa	gatgccattc	ttgggtggcct	ctgcacccaa	attgcatctg	gaaagaacta	600
gggtctcatt	cagaatgtcc	aaaaggaaat	tcttaagagc	ttaaattcag	atttgtgtct	660
cattaatgca	gtgaacaatt	caaaaccaca	cagattcctt	ggcaggaagg	ataatggaat	720
aacagtgttg	atgagacctt	tttagcttca	aggtttcggg	gtctaaacaa	atggatgatt	780
catttggaat	gaaactcaca	atgcaagtag	aaggacctct	ccaaatcagg	ccagtggggt	840
tatcctgggt	tggaaatctg	tgtgaaacca	taggtcttaa	cactctggag	cagcacattg	900
ctgtggatat	gtccaggaga	ccttagatat	ggcttaaagg	ctttcaagat	gaggacagaa	960
attgcttaca	attgctcagt	ttctcaacag	aaagactcat	aagagtgcc	gcatggggta	1020
catggagtga	agctgggtgg	gaagcatcat	ctgcacagtc	cctgtcctag	tgaggacttt	1080
ttctctgtat	gttttcatac	catgggattt	ttggatatca	gtgtattttg	gttcttgaaa	1140
tagcctaata	gctgctcaca	cattgggtag	gaatattata	ccaatgtcat	ccccaaagga	1200
aggggtgagct	gaatggaaat	taagcccagt	cattttattt	gatctattag	ctctgttata	1260
agtgcagatg	caccagatc	accctcctca	gccacacag	tgctgaacca	tcttccctcc	1320
tgttctccat	ggctattaat	agtatagcta	aatttagagt	gcagagccag	atataagtat	1380
tttggaatta	tctcccagtt	tgtggtagaa	gctgactgga	atcacagggt	agtatctctt	1440
atccaaaatg	ctagggacca	gaaaggtttc	agattttttc	agatttttga	atacttaaca	1500
gttgagcacc	ccaaatctga	aaggcttctg	aacgtcatgt	cagcactcaa	aaaagtggat	1560
tttgaggcac	ttcaaatttc	ggatttttgg	atttgggatg	ctcatcctgt	gtaggagagg	1620
ctactcgatt	ccatttaatg	actgtcctag	tcataatcat	ccaaagataa	aagccaggta	1680
gatgttgaaa	gctctttcca	gggtgaaaa	agtgttctta	cgttctctgc	atgtgactag	1740
catcactgtg	gaaattaatg	ctctgttctt	cactagaatg	tagtaagtgg	ttaaactgag	1800
ctatcccca	cctgatgact	attggcatcc	atttgcaagg	ccaatggcct	ggattaaggg	1860
ttaggattat	ttgtagctag	aaggtaattt	tatttctgtg	aaactaattg	gctcatattt	1920
gagggttaggt	gtggccttga	ccttaccagt	acattttatac	ccactaccag	ttgactagcc	1980
cagataattg	ttaaatgggt	cttcttttct	gcttctcagt	agacttccat	gccattacaa	2040
aggaaatttg	aattacctag	tgtttgtata	ttccatgata	actatgtata	acttctgtta	2100
cacagcttat	gtattgttaa	catttaagtg	taaaccatgc	cacagctaac	acttaaaaa	2160
gaaaactaat	tagttcttgc	ttagggaaaa	tgccaggat	gaagtatggc	atatacttga	2220
cactgtcctg	tgtaacctt	tactttgctc	aggctttcaa	gattgagtct	tttttcccc	2280
aaattaggtt	aacatgcatt	tgaccccaac	ctgtgggggt	tgagtaagct	ggaaatctgt	2340
gacggtaggc	tttctagtgt	cacgagggtg	tggtgactga	aggaaaagct	gggatcacag	2400
gttcttctg	atggagagga	aggtttattt	ctatgcccct	cccaccaccc	tccacctaga	2460
gctcacccaa	gcctgtctca	gtcccagggg	caggccattc	tgcaaaagca	ggacctcaca	2520
gaaacaaggg	ctgggttgag	gtcacccctt	tcagagttgg	ttcctggcca	gatgggtaag	2580
aggcatttgt	aattttaaaa	atgtgaaact	tgggttttgt	gttttcttct	aagtgcctaa	2640
ataagcaagc	caggctgttg	atatttttagc	cagagaaatc	ggcaagccaa	gattaaccgc	2700
aatctgaagt	ttagaatctt	gagtttgcat	ctgcatcata	tcagtctgtt	ttgatgagga	2760
aacatttgcc	actgaggagt	tggagggagg	gcaagacgac	agtgttaagt	cagatcattt	2820
aatgggttcc	cctaagccct	ggaaaaatat	ttgaaagaat	ggcagcaaaa	agggttaagaa	2880
agcaagccag	atttactgca	caatatgcag	taccagctac	tacttttaat	cccaagagaa	2940
cagtgtgatg	tctaataatat	acaggtctat	gaaaatactg	tggaataagc	ccaggaaggt	3000
tagatgtgtt	tgcaaaataag	ttgccccaaag	ggtccccctc	taagtaaaac	aaatattcag	3060
accacaggct	ttaatgtaaa	ctgtcaaaaa	gtgggatgtg	gaggattttt	gttaagtgtc	3120
aatcgaagtt	aaaaagcaag	ggtttttggc	caggcgtggt	gtcacgcct	gtaatcccag	3180
cactttggga	ggccgaggcc	ggcaaatcac	ctaaggctcag	gagttcgaga	ccagcctggc	3240
caacatgggt	aaaccccgct	tctactaaaa	atacaaaaaa	attagcccg	tgtgggtggca	3300

30

agtgcctgta	gtcccagcta	cttgggagggc	tgaggcagga	gaactgcttg	aacccgggag	3360
gtagagggtg	cagtgaacca	cgatcatgcc	actgcactcc	agcctgggca	acagagcaag	3420
actccatctc	aaaaaaaaaa	aaaaaaaaagg	cggccgctct			3460

&lt;210&gt; 45

&lt;211&gt; 2622

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 45

aggttcccctc	ctttggatat	ttggcagttg	tacttttatgc	agttcagcca	agaactaaac	60
acaggccttc	agggtgtggc	agagtacagt	gggccagagg	tcctttcctt	tcttcagtgtg	120
gtcatacacc	tctgttaatt	caacttaagt	ttatattact	tttgggggtg	ggagtgggga	180
agaggatcac	attgttgact	tgctggaagc	attatgtcaa	ctaaaatcct	tcagttctta	240
ttttcatcat	gctgttgggt	tccccctatg	ttgtttcctt	ttaaaaacac	caaatgcaga	300
acttcccttt	tatactgctt	ctatttcatc	ttgttgactt	gtagcctatc	agagtatgat	360
tcctttattgt	catctaaggt	attctgatat	aatcactgca	tgtctaaatt	ctagtatttg	420
actaaatgct	acttgagacc	attctccttg	ttgtcacatt	atctttacagt	tgaagaaact	480
gaaccttaga	gagggttaagt	accttggtta	aggccacctg	gctggaacat	aacatcctgg	540
tcatttttaac	tactaatgct	gtcatgaaag	ggagttagcc	agggcaggtg	tgaatagagc	600
attctagcat	aagcagccaa	caattagagc	aactgtgtat	ttttagtgtg	tggatagcaa	660
atctttgctg	tcttggaagt	tactatctta	gaagcatggg	attaagctct	ctattcaatg	720
tggtgttcac	atcagtagca	ttggtttcat	ttgggagctt	tttagaaatg	cagagctcag	780
gccccactcc	aggcctaagg	attcaatctg	cattttaaga	agtgacctat	atacacatta	840
aagggtttaag	aagcattgct	gtagtatcct	atctttccaca	ctttagtcat	ttgcatatca	900
cttgtgcaac	tttggtctta	tctgcatact	tcctgaaggg	ggcctgcccc	tcacacctg	960
tgggtatttc	tcgtcaggtg	gagatgaaag	aatgagaaaa	gaaataagac	acaaagtata	1020
gagaaagaac	agtggggcca	ggggaccggc	acactcagca	tgcgaggacc	tgcaccagtg	1080
ccggtctctg	agttccctca	gtattttattg	atcattatct	ttactgtctt	ggcgaggcga	1140
gtgtagcagg	gcaacagggt	gcgagaaggt	cagcagggaa	acgtgagcaa	aggaatctgt	1200
atcatgaata	agttcaagga	aaggtaactgt	gcctggatgt	gcacataggc	tagatttatg	1260
tttcacttta	cacaaatata	taactagcag	agagcaacaa	agcagtattg	ctgccagcat	1320
atctcgctc	cagccacagg	gcgggtttct	cctatctcag	aatagaacga	atgggaatgg	1380
tcgactttac	actaagacat	tcatttccca	gggacgagca	ggagacagaa	gccttcctct	1440
tatctcaact	gcaaagaggc	ctccctcttt	cactactcct	cctcagcaca	gaccttttac	1500
atgtgtcggg	ctgggggatg	taagggtctt	cctttcccac	gaggccatat	ctcaggctgt	1560
ctcagttggg	ggaaacctgg	acaataccca	ggctttcttg	ggcaggggtc	cctgcggcct	1620
tccgcagtgc	attgtgtctc	tggttaatcg	agaatggaga	atggcgatgg	cttttaccaa	1680
gcatactgcc	tgcaaacata	ttgttaccac	ggcacatcct	gcacagccct	aatccatta	1740
aaccttgatt	caatatagca	catgtttctg	ggggcacaga	gttggggcta	aatttacaga	1800
ttacacgcat	ctcaaagcag	aacaattttt	cttagtacag	atcaaaatgg	agtttcttat	1860
gtcttccttt	tctacgtaga	cacaggaaca	atctgatctc	ttttcccccac	acttcttata	1920
tggttaatgta	ttttaattgt	ctgttagctt	ttatgtcttt	ccaacttgaa	ataatttcac	1980
ttacaaaata	gctggaaaaa	tagtaaagat	aattcctgta	ttttcttcac	ccacgtatc	2040
ccaaatgtgt	aacacctaca	atcagcaaaa	atcagcacat	taatattatt	acctaattcta	2100
cagacctcat	tcacattttg	ccagttgttc	tgctggtgtg	ctgtatgtag	tgcaggatcc	2160
agtctaggcc	cgtatgttgc	atcttgctgt	catgtctctt	tagtctcctt	tcctctcagt	2220
ggctctttct	ttccatgac	tatgacagaa	atgatttgta	tccttcata	atctgcagtc	2280
acatgaagtc	tatttgtccc	actactggtg	acgtaaaatt	tggttaactg	gttgaggggtg	2340
gtaccgtcta	ggtttctcca	tattaagttg	ctatttttcc	ttgtgtatct	aataagtatc	2400
ttatagggaa	atcctttgag	actatgtaaa	tatcatatta	catgtatttt	gcctatttta	2460
gcactctttg	aagataattg	cctgaaacat	tttttactgt	gatgatcttc	taattctatc	2520
attctgcatt	tattagtttg	gaattcgatt	gtaagggaag	atattccctt	gtttcctatt	2580
atctttattaa	agaaatttac	atccaaaaaa	aaaaaaaaaa	aa		2622

&lt;210&gt; 46

&lt;211&gt; 1984

31

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (106)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 46

```

agttgttgct atgaatgatt ttaagccttt ttttttaagc ttggcaaaca tcccagctaa      60
tcaaaatagt catattcctg agaagtagga aactaaaact tctttncata taattgtag      120
aagggttggt tcccaacta ccatagttac aaagggtgaat agccaaatgt taggaacaga      180
atcaaaaagaa taaaaatctg tgaagagatc ctactactct tccttctatg ttttggtttt      240
ggtttctatt gtccctatca tttcagcaag tggaaacagca gcaagttttt cagtgcata      300
gctgcacaag aacaaaatat aaatctgtat ggcaacaaaa atcaaagtga aaaccaaaacc      360
aaaaacccaa acaccctatg taactatcgg aggcataatac gtggtataaa tgactgtagc      420
tgtgatacac acatggctac ttgtcacatc actttccata attatttact gcaaaatgat      480
tgagaggctt ttggtgcagg cagccgttaa cctctgctt cctttgttac ctctggatta      540
ctttgcagta aattgcagggt cttttaagag atttaagctt cagttttctc aaaccaaaacc      600
aattatcctg tcttatctga agatgcagggt ttgtgggcaa aagaggctgg ttataataat      660
gccctcatat tgagtggctc gtaaacggct gcacacttca ggcactgtag ttgctgaaga      720
tgctttgtta aatgtgacct tgactggctt tacaggggtg tagaatgtaa tctacacaag      780
gtgactttgc atctatcttg ctcttgaggt ggatgaaatt gagaagctgg agtgtgtaag      840
ccatgcacat aagtattctt cactgtaaat tttgttttca tttttaaccc aattatggta      900
ctttatccaa tgcacaactg atctctcagt agatattcat ttgaaaatag tgtggccttg      960
atcagtgaga aaggaagga gaaaagtgac ttttttgctt atgtagaaat gactcatttg      1020
ctgagagttt gtctttctgc agcactcttg gtataatgtt agtgatcggg ctcccttttg      1080
attggggaaa gttaattgtt ttgaccctgg agttaattca gttgagttat cttatatttt      1140
taggaagtat cagaattgct ctgatgaata acaaagttga ctgttttgat gtccaatctc      1200
aggttttaga atatagtggt gtaaaagtcac actattttta attcttaaaa caactttaat      1260
ttcgtacacc ctaaaagtca catgcataag gcctgttcag agagcagagc ctccatcttt      1320
ttgctccttt tctactttgt acttcacttg raaaaatc aagtgacttt acatkgata      1380
tttccattgt aaccttgaca tttctcaaag ataaagcact ttttgatcat gaaatacatg      1440
aaatctttgt gtgatgtgga tcatagtttc tcaggctccc ttagataatt gcttatgaat      1500
attgttctaa ctctgtgtaa gaagagtaga aatctttgct aatgtagtaa gggttgatt      1560
attgatccag aatgcatttt gctagtttcc aatggatggg agagtaaata atgctgcatt      1620
cacaatttaa taagttactt tcccttgagc cttaaggtaa ctttttcttt tctgtcaact      1680
acagcactga agttatagta agtgaatgag attatcagtt ttcagggttg gtttttagagt      1740
actgtaaatc aattagctgt cttcctaaag agttacaact cccattcagt atactggata      1800
atgggtgtgt ggggtggggt ggggagggcg ggagatagtt ttagaaaaag aaaaaagaaa      1860
aaaaaaaaaa acacatttac ctttaagaaa tacagacaaa aaaaaaaaaa aaggcgccg      1920
gctctagagg atccaagctt acgtacgcgt gcatgcgacg tcatagctct tctatagggt      1980
cacc

```

&lt;210&gt; 47

&lt;211&gt; 1987

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (442)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (444)

&lt;223&gt; n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (473)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (493)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1011)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1025)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1111)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1119)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1169)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1234)  
 <223> n equals a,t,g, or c

<400> 47  
 ggcacgagac cggcgctagt cttgcgtcgc acagcaggat tgcgctgctt tctgatgacc 60  
 ttcgacatgg atggcgatgg taatgaacgc ttttcccgtg acgccccacc cgcaaccgac 120  
 ccaggccgga atgaccctga tgcaggtcct ggtatctgtg ctgatccctg ctgtcggcct 180  
 gctaagggcg gcagtcattc agctcaatgc attgaaatac accgacagtt ccaggatgac 240  
 cagtcaggcc agtttcattg cctacgacat gctcgaccgg atccgcgcca attcgggtgc 300  
 tgattactcc tggggccagg gtgaacgcgc gccctccacc acctcggctg cgagtgtgcg 360  
 tgatctggac ctgcacgact ttgaagcgaa tatcgtcggg ttcgcccggg aaagcgccaa 420  
 ggggtccggt gcggtcaatc ancnaagaagt gaccatcagc atcagttggg acnactcccc 480  
 tggagcgaat gcncaaggca cccgggaaac attcaccctg accagccggg ttgcagtcga 540  
 tccgaggggtg ttgccatgag aggtccagtg tgcggtttca gcctggtgga gatgctgctg 600  
 gcattggccc tcggcctgat gttgatcctg ggggtgaccc aaattgcact cagctcccga 660  
 accacttatg ccagccagag tgcggcttcg ctgttgacag atgatgcacg gttcgcctt 720  
 ggcaagctga ttcaggaaat acgccaggcg ggcattgttg gctgcttgtc cgctgcatca 780  
 atcagcaacg ctcccgcagg ttttgatcgt cccattggat ggagtaccac cggcagttcc 840  
 cggtccttga cgctggtgac cgccgacgtc ggggaggggtg gcagcaagcc ggactggacg 900  
 gtgctttccg attgcaccgg ctctgcccac gcctatgttg gaagcccggc ggcagcgaac 960  
 gcccgggcaa atccacttcc cacttgcgca aagctgacct aacacctttg nagggcgggc 1020

33

aagcnggaag	ttaagtaacg	ctggcggccc	ccgagcaaag	cggtggttgg	tggataacgt	1080
ggggcattcg	atatcagttt	tcggcgtggc	ngacaagcnt	ggctcaacgg	ttgtcagccg	1140
atatgacccc	acccccggcg	atgagtcgnt	catccgcagc	gtgcggattc	tgctgacact	1200
tcaggatcca	aatgggttgg	tgaagacca	ggcntacagc	gtggtcgcgg	cactacgtaa	1260
tcgcctggag	tagcgtgccc	atgggttatt	acctctcccc	ttcgaggcag	gcaggcatgg	1320
ttttgctgat	cagcctggta	ttcctgctgc	tgttggcact	cctcggagtg	tcttcgatgc	1380
agggagcaat	ctcgcaagaa	aaaattaccg	gcagccttcg	gcagcgcaac	cagtcgtttc	1440
agcaggccga	aagcggcctc	aggcttggcg	agtctttggg	gcaggcgctca	ggtttcgccc	1500
tgcgcccttg	ccactcgacg	gctgcgtgcg	cgcacactgc	cgaatcgggt	tcggtagtgg	1560
ggccggggac	gaaccccgta	tcgactgtga	cctggatagg	gatgaaagat	ggcgtctacg	1620
gtattcaaaa	cctggggccg	ggaacggggt	tgggtcaactc	ccggcagagg	cccaggccac	1680
ggctctatcg	gtgacatcag	tggcgctcag	tgggcactcg	cgttcgggtcc	tggagtctgt	1740
gtatgcccgt	gtgggcagcg	ggccggcgga	gcgtttccga	cgaatcatgt	ggcgacaact	1800
tcaataggtg	agcagcacga	tgggcaagga	ttgcacaggc	ttcaccctga	tcgaataact	1860
gatcgccgtg	gcctcgtgcc	gaattcggca	cgagattaat	tccccataaa	atctttgaaa	1920
tagggcccgt	atttacccta	tagcaccccc	tctaccccct	ctagagccaa	aaaaaaaaaa	1980
aaaaaaa						1987

&lt;210&gt; 48

&lt;211&gt; 2113

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 48

ggggaatatg	gatggctttc	ttctttgccc	tctttgttat	tttctttggt	attgttgttc	60
agatggagtc	tcactccggc	ctgggcaaga	agagcaaaat	tctgtctggg	gggcaggggg	120
aagagggtata	ttttttggat	taaaatggaa	tgctttccctc	aaactgttaa	aattaggtag	180
tgttacaagt	ttaaagaata	ggttttagcca	gtgtgattat	tgcagtgtaa	taaaaaggat	240
cccagggtggc	acacatgcta	gtgtgggtgcc	cagatgtcca	actgaagtga	aggataacat	300
gcgtttcaac	attttctata	tactaggcag	tttttcttaa	gtcttggttt	attgtgatat	360
aatacagtaa	aatttactct	ttatacgtat	gattccatta	atattaacaa	atttaaacag	420
tcttacagct	aaatacaatc	aagatacaga	acatttccat	cactacagga	agttgttttt	480
tgctaattta	actaaacttc	ttaccatgct	caacccccca	ccccaccct	gagaaccact	540
gctttgtttt	tctgatcata	tagtttggtc	ttttcaaawa	tgtcatacag	ttggaatcat	600
atagtatgca	tgaatcatgt	agctgaagtt	ttcatttcac	ttaggtaaac	acctaggaat	660
aggagtgtcg	ggttatactc	taagtgttaa	acttcataag	aagctgctaa	actgttttcc	720
aaagcatctg	taccattttc	ctttcccacc	agcaataaag	cattcattag	tctacgtact	780
caccagtgtc	agtgtggtca	gaatgtatag	ttttaattat	acacattata	atagatgtag	840
agtgggtatct	catcgtgatt	ttgccttttc	ctaataaata	tctttcctta	catatttgtc	900
attagtgtgt	ctttctttagt	aaatggtcct	attgttttgc	ccatttttaa	aagttgagtt	960
ttcatattgk	tcaattttcga	gagttcttca	tatatcttgg	acacaagtcc	tttgtcagac	1020
atgtgatttg	caaataattt	ccccagagt	ttttcctgtc	ttgtcagtc	cttaatagtg	1080
tttttaggaa	ggagaacagc	tgcttttaat	tttgataaag	cttgcaattg	tttttctttt	1140
atggatcatg	cttttggtgt	agtatctgag	aactccttct	taaaccagtc	acacatgtct	1200
tctgttttct	ttgaagcttt	acaggttgag	gtacattttg	gtccatgatc	acttttgagc	1260
tactttttat	atataggtag	tataagatat	ggtttgaggt	ttttgttttg	ggtttttttt	1320
cttcctgcat	atgaatgttc	agttgttcca	gcaccatttg	ttgaaaaaac	actatccttt	1380
ctctagtaat	ttgctttttt	acttttgta	gttgactatt	tgtgggtctg	tttttaaact	1440
ccattttata	ttatgtctac	attatgaact	ttatagttag	tcttgaaatt	mgtaatgtg	1500
ggctcttctca	cttgtgcttt	tcaaaattgt	tttggctctt	ctaattcttt	cactttttcc	1560
atataaattt	tagaaamagc	ttattgattt	ctaccacttt	cccccaaaaa	gccacttggtg	1620
aatttgacta	agtttacatc	gaatctaata	aatgattttg	gagagaagtg	gtatcttagc	1680
aatacagtct	tttctgataa	cacttcctgt	tttagttctt	tttctcattt	aataattttc	1740
caacattaaa	accttggaac	aatttagatg	tgtatctgaa	taattccagt	tttgttgcca	1800
tttaaaatgg	tacataatcc	ttattttata	gatgagaaaa	ttaagtaact	tgcccaagtc	1860
acacagttac	taaatgacaa	agctagattg	aaatctatga	gcttataaac	taagtgcctg	1920
ttttgagcta	cagtactagt	tttattaaat	tgtgtggtt	aaactgaggt	gggaggatag	1980
tttgagcacg	gaagattgag	gctgcagtga	tctttgatcc	tgccattgta	ctccagcctg	2040



34

gcaacagagt gagaccctgt ctccggaaaa aaaaaaaaaa caccatggaa gcaagcaaaa 2100  
 aaaaaaaaaa aag 2113

&lt;210&gt; 49

&lt;211&gt; 3465

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 49

agaaaaatgaa	agcaccaaag	agccttcttt	gctacaatat	ctttgtgtgc	agtctcctgc	60
aggattaaat	ggtttcaatg	tacttttatc	tggcagtc	accccccta	ctgtgggccc	120
gtcctcaggt	cagctgccgt	ctttcagtg	cccttgc	gtcttaccat	ctccacctct	180
gggccctttt	cctgttctct	attctcctgc	aatgccgggc	ccggtttctt	ctactcttgg	240
tgctctccca	aacacaggac	ctgtgaattt	cagcttgcc	ggccttggat	caatagccca	300
gcttctcgtc	ggccccacag	ctgtgggtta	tccaaagtcg	tccacactcc	cttctgcaga	360
ccctcagctt	cagagtcagc	cctcactaaa	cctaagtc	gtgatgtcaa	ggtcacacag	420
tgctcgtccaa	caacctgagt	cccccgttta	cgtgggacat	ccagtctcag	tagtaaaatt	480
acatcagtc	ccagttccag	tgacccccaa	gagcatccaa	cgcacacatc	gtgagacgtt	540
tttcaagaca	cccggcagcc	ttggagaccc	tgctcctgaag	agaagagaaa	ggaacaatca	600
cgaaacacca	gctcggcccc	gaggagacta	gaaatcccca	gcggcggcgc	tgactaacct	660
gccgctttgc	caggtggggg	tgggatcaaa	cgccctgaga	gtcccggatg	tccgaggcgg	720
gatgcaaacc	atcccgtcct	gagcacgggt	ccttctcttc	tctttcatcc	acacttctgt	780
taacttccca	ccaccatcaa	tcatctgatt	tcctgaaagt	aattaattgt	gcatttaata	840
ccagttagag	ttccgactct	gcatgggtgc	acagtgaag	cgccgactga	cttatggttt	900
tgattcaaga	atcgtcttat	tgctggaagt	agatctgaat	aggataccgg	agccttgttt	960
ttctaaaggg	ggcgctgtc	tagcacttaa	ctagggtaag	cattcttaac	atgtatttcc	1020
acttgccttg	agtaaatctg	tggtgagaga	agcttctctt	ctgcagttta	aaaaagctac	1080
tgcttctcta	ggcttcatca	ggaagccatc	ttcagttgtg	aatcctatgg	tgttatttat	1140
tttgttcctg	aaatgggatt	tagtgcaaaa	agtttacaac	tacagtcttt	aacacatttt	1200
tttcagggta	tgacgacttg	aatgtttata	cttttattct	ataatttgcc	ctgcacttat	1260
tttacaacct	agtaataatg	tggataaatg	tatctacatg	acacatgtca	agacccaaat	1320
aactgtgaat	gacacacctt	gctgtaaatg	aactgtgcta	accctgactg	tgggcttgag	1380
aacaaagatg	aactctagaa	ctctagcagc	ctaactgctg	cttctcaaat	aactgtgtga	1440
acagtgagat	attactgttt	gtttctaaaa	atcctactgt	gcccagtttc	cttactaca	1500
tgccctgcat	tttttattta	aatatttagc	tgtagcgcca	tcagatatgg	atgccttcta	1560
acaattgctg	tttgtaaaat	aaatcaggat	ggtagaaaag	gattttatgg	aaaattggaa	1620
cctggatgag	accttttctg	tgaattctga	agagtaatga	tgcgaaaatt	gatacagggc	1680
aagagatgat	ctttttgttt	ttcttctact	tcatgtccag	aagagtaaga	gggaaaatgg	1740
acatatgttt	catatccaag	ggatttcaaa	ctgtagttag	ttggtacctc	tgaaaaatga	1800
gaatgggtgag	cgcacggggt	gggtgttcta	gcataaatac	aattctggaa	actgttatgc	1860
aatttccctt	ttttaaccca	cattacttta	ggggtgcatt	aagtcgccaa	actatactag	1920
ttctttgtat	tcctagactt	gctgatattt	acctctctct	tgtctcttca	gagtaaatgg	1980
ttcccttctt	tccttctctc	tttcttctat	tctctcttcc	ttccctctct	cctacttctt	2040
ttcttctctc	ctcttctctc	cttaaaacta	tcttagatgt	agaatcctgg	tgtagggttt	2100
tatttttatt	ttattttttg	acccaataaa	atgttatatg	aaagaatgaa	aatattaatt	2160
taagagactc	tgggagtctg	aataaagtag	ctttatatta	actacaggat	aatattagcc	2220
ttattacccc	cacaagattt	tttaaaactt	gaggtaggta	gctacattaa	ataaatttgc	2280
tacttatata	aaaattttta	tcaacactaa	acttttaaag	tttacaagtt	ttttttttct	2340
tttttacagt	cttctataga	gttaggttaa	aaatgtgggt	ctaaccatca	acaattgc	2400
ggttaaatga	ccctgaacta	aaactgatgg	gttccctatc	aaaacaaata	aaaataatac	2460
tttttcaggt	ttcaatctgt	gcagggtata	tgcatgttaa	ttctaccatg	cttaagaact	2520
tccacaaaat	atttcatgga	gagggtctga	tttagacgga	aacagaaatt	gcttttcccc	2580
tcactgttcc	tgaatgctct	atacttgttt	taacattttt	gctatctttt	tttattattc	2640
tgatcatgat	atgaccattt	aacctcagaa	ttcataattc	ctgaggggtg	ttaagaagca	2700
gtcccattgg	tgaggatatt	atgacttgg	gaccattctt	aggagtagaa	aaccaaggac	2760
aattgcttct	gtattcagta	tccacttctt	aatgtggctt	tatatgtaaa	aataataatg	2820
cagtgggtgt	ttctgtcagg	aaaataaatc	ttacagaaca	actgggtgaa	ttgaagctgc	2880
tgcgctagac	ttggatattt	tgggtagtga	agaagcaatg	gcaatcttga	gtctattatt	2940

35

gtataattta	gtaaaagaaa	aaaataatcg	ttgggtggcc	tactaagaga	atgcagcttt	3000
tttgagttgt	cacagaggct	gtgtgtgccc	tacactgacc	agggtttgta	aaaccctttc	3060
attctgttac	aagagtcggg	ggtataactt	ttataactga	atctacctac	caagttttaca	3120
tttctcaatt	ccttttttga	aggtgctatt	tctgtattta	aataactttc	ttttaacgta	3180
aagctgcttt	ctgcttatct	tattgcaactg	ctagttgtat	gtaggtatta	attttattgc	3240
tgcttactgc	ttttgttttc	ttattattta	gctctgctct	ttttccta	ggctatatta	3300
tctatagcta	tttacttgta	actgtactac	atgtaaactg	attttttgtt	ctgatttttt	3360
ttctaataatt	tttaggaaaa	tattaagctt	tataaaatag	caataaaaaa	taattcattt	3420
aaaaaaaaaa	aaaaaaaaact	cgtagggggg	gcccgtagcc	aattc		3465

&lt;210&gt; 50

&lt;211&gt; 1237

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 50

atggggccagg	aatgggggtcc	cggggcatgg	tgctgggcct	cctgggtgcag	atctggggccc	60
tgcaagaagc	ctcaagcctg	agcgtgcagc	agggggccaa	cttgctgcag	gtgaggcagg	120
gcagtcaggc	gaccctggtc	tgccagggtg	accaggccac	agcctgggaa	cggtctccgtg	180
ttaagtggac	aaaggatggg	gccatcctgt	gtcaaccgta	catcaccaac	ggcagcctca	240
gcctgggggt	ctgcggggccc	cagggacggc	tctcctggca	ggcaccagc	catctcacc	300
tgcaagctga	ccctgtgagc	ctcaaccaca	gcggggcgta	cgtgtgctgg	gcggccgtag	360
agattcctga	gttgaggagg	gctgagggca	acataacaag	gctctttgtg	gaccagatg	420
acccacacac	gaacagaaac	cggatcgcaa	gcttcccagg	attcctcttc	gtgctgctgg	480
gggtggggaag	catgggtgtg	gctgcgatcg	tggtgggtgc	ctgggttctgg	ggccgcccga	540
gctgccagca	aagggactca	ggtaacagcc	caggaaatgc	attctacagc	aacgtcctat	600
accggccccc	ggggggcccca	aagaagagtg	aggactgctc	tggagagggg	aaggaccaga	660
ggggccagag	catttattca	acctccttcc	cgcaaccggc	cccccgccag	ccgcacctgg	720
cgtaagagcc	ctgccccagc	ccgagaccct	gccccagccc	caggcccggc	caccccgctc	780
ctatgggtcag	ggtctctcct	agaccaagcc	ccaccagca	gccgaggcca	aaagggttcc	840
ccaaagtggg	agaggagtga	gagatcccag	gagacctcaa	caggacccca	cccataggtta	900
cacacaaaaa	aggggggatc	gaggccagac	acgggtggctc	acgcctgtaa	tcccagcagt	960
ttgggaagcc	gaggcggtg	gaacacttga	ggtcaggggg	ttgagaccag	cctggcttga	1020
acctgggagg	cggagggtgc	agtgagccga	gattgcgcca	ctgcactcca	gcctgggcga	1080
cagagtgaga	ctccgtctca	aaaaaaacaa	aaagcaggag	gattgggagc	ctgtcagccc	1140
catcctgaga	ccccgtctc	atttctgtaa	tgatggatct	cgctcccact	ttcccccaag	1200
aacctaataa	aggcttgtga	agaaaaaaaa	aaaaaaa			1237

&lt;210&gt; 51

&lt;211&gt; 1397

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1383)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1396)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 51

ggctccctagg	agttgagcag	gaacaggcat	ctgtggttta	cggcgacctg	gctctccgcr	60
ggccacgtgg	gtgggtgagg	cacacgagtg	ggaagcgcca	ccgacgtgtt	tctccccgac	120
cgtggctttg	ccaaagactt	ttaatagcat	tttttaagtg	caaaacgtct	aggtaaaaa	180

36

ctttatcatc	agtgacaaaa	ttagaatgta	tttaatatag	taggtgggtt	aagaactggt	240
ttaacgtaag	acaaactgat	agcaacattc	tgttgtttta	aaggaagtgg	gtccgtgaca	300
ttctgcagct	agtcactac	tccaaggtaa	ctatcgactt	ggtttcagt	aatctatttt	360
gtttttaact	acagtgattt	attagctcag	tatctagaaa	ttacgtatat	tttgtgtac	420
tgtcatcgat	gtgtaaaactc	tgtttttatt	tgtatttatg	cacttggttc	ccatttggag	480
cctctgggtc	tttctgggat	aagtgggtgc	tgccgagaca	tctcccgtt	gtcagtggc	540
aggagcagct	gagctctagt	ctgccagctg	ctctgctctt	tctgggaagg	aggtggcgcc	600
cgccctcag	ggtgtctcca	gggctcagct	tccgggtgg	tagagctggg	gagccccagg	660
ggtgggggga	cagctgggag	atggaggtgg	cacctgctcc	cctagatcag	tactggctct	720
gaggacaggt	gagcagtggg	aagaccaaag	aatggctggc	agcgtgcca	rggttgga	780
tgggggcaag	atcctggggc	tgtgtgccct	ggggcctccc	tcacctgtct	tgggtggccat	840
ggcctcaggg	atggctccta	ggtggctgag	gcacagcagt	ggctggaagg	tgccccgtgg	900
aggctgaggt	ggaggcgcg	ccagcagctc	ccccctgtgg	ccatggcggg	cacgggscgt	960
aggagctggc	tggcgccgg	ctctgcatgt	tctgttgcc	tgtcgtctgt	aactctagt	1020
ttcgacattc	gccgtgatac	agtgggtgtca	cgacgtgtgt	aactgtggc	agcagacctt	1080
gttccgcgtg	gacgcctcaa	gtggattaat	ttctggaagc	ctcaatctgt	atgtttgagt	1140
atttacatga	gaatgttatt	tgaatggaat	tttcttaacc	cagaaggtag	tatttataat	1200
catttacttg	tagcgaactg	tttaaagtta	acacttggtt	aaattttttt	acactatagc	1260
atttatgcaa	tggtttacag	aattcatgga	gttattttta	tcagtatggg	aattaattaa	1320
aaccttgaat	cttaaaaaaa	aaaaaaaaag	ggcgccgct	ctagaggatc	caagcttacg	1380
tangcgtgca	tgcgana					1397

&lt;210&gt; 52

&lt;211&gt; 2271

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 52

cggcagcagc	ggcagcagta	tcggagggct	ttggaccatg	aggaggaggc	cctgtcatcg	60
ggcagtggtg	aagaggcaga	agccatgtta	gatgagcctc	aggaacaagc	ggagggtccc	120
ctgactgtgt	acgtgatatc	tgaacactcc	tcacttcttc	cccaggacat	gatgagctac	180
attgggccc	agaggacagc	agtgggtgcg	gggataatgc	accgggaggc	ctttaacatc	240
attggccgcc	gcatagtcca	ggtggcccag	gccatgtctt	tgactgagga	tgtgcttgtc	300
gctgctctgg	ctgaccacct	tccagaggac	aagtggagcg	ctgagaagag	gcggcctctc	360
aagtccagct	tgggctatga	gatcaccttc	agtttactca	accagagccc	caagtcccat	420
gatgtctact	gggacattga	gggggctgtc	cggcgctatg	tgcaaccttt	cctgaatgcc	480
ctcgggtgcc	ctggcaactt	ctctgtggac	tctcagattc	tttactatgc	aatgttgggg	540
gtgaatcccc	gctttgactc	agcttctctc	agctactatt	tggacatgca	cagcctcccc	600
catgtcatca	accagtgga	gtcccggctg	ggatccagt	ctgcctcctt	gtaccctgtg	660
ctcaactttc	tactctacgt	gcctgagctt	gcacactcac	cgctgtacat	tcaggacaag	720
gatggcgctc	cagtggccac	caatgccttc	catagtcccc	gctgggggtg	cattatggta	780
tataatgttg	actccaaaac	ctataatgcc	tcagtgtctg	cagtgagagt	cgagggtggac	840
atgggtgcgag	tgatggaggt	gttctctggc	cagttgcggt	tgtctcttgg	gattgtctcag	900
ccccagctgc	ctccaaaatg	cctgctttca	gggcctacga	gtgaagggtc	aatgacctgg	960
gagctagacc	ggctgctctg	ggctcgggtca	gtggagaacc	tggccacagc	caccaccacc	1020
cttacctccc	tggcgagct	tctgggcaag	atcagcaaca	ttgtcattaa	ggacgacgtg	1080
gcatctgagg	tgtacaaggc	tgtagctgcc	gtccagaagt	cggcagaaga	gttggcgctc	1140
gggcacctgg	catctgcctt	tgtcgccagc	caggaagctg	tgacatcctc	tgagcttgcc	1200
ttctttgacc	cgctactcct	ccacctcctt	tatttccttg	atgaccagaa	gtttgccaac	1260
tacatcccac	tcttctctgc	tatggctgtg	cccatcctcc	tgtccctggg	caagatcttc	1320
ctggagaccc	gcaagtccctg	gagaaagcct	gagaagacag	actgagcagg	gcagcacctc	1380
cataggaagc	cttcttttct	ggccaagggtg	ggcggtgtta	gattgtgagg	cacgtacatg	1440
gggcctgccc	gaatgactta	aatatattgtc	tccagtctcc	actgttggct	ctccagcaac	1500
caaagtacaa	cactccaaga	tgggttcac	ttttcttcct	ttccattcca	cctggctcaa	1560
tcctcctcca	ccaccagggg	cctcaaaagg	cacatcatcc	gggtctcctt	atcttgtttg	1620
ataaggctgc	tgcctgtctc	cctctgtggc	aaggactgtt	tgttcttttg	ccccatttct	1680
caacatagca	cacttgtgca	ctgagaggag	ggagcattat	gggaaagtcc	ctgccttcca	1740
cacctctctc	tagtccctgt	gggacagccc	tagccctcgc	tgtcatgaag	gggccaggca	1800

37

ttggtcacct	gtgggacctt	ctccctcact	cccctccctc	ctagttggct	ttgtctgtca	1860
ggtgcagtct	ggcgggagtc	caggaggcag	cagctcagga	catggtgctg	tgtgtgtgtg	1920
tgtgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	agaggttcca	gaaagttcca	gatttggaa	1980
caaacagtcc	tgaattcaaa	tccttggttt	tgcacttatt	gtctggagag	ctttggataa	2040
ggtattgaat	ctctctgagc	ctcagttttt	catttgttca	aatggcactg	atgatgtctc	2100
ccttacaaga	tgggtgtgag	gagtaaattg	gatcagcatg	taaagtgtct	ggcgtgtagt	2160
aggctcttaa	taaacactgg	ctgaatatga	attggaatga	tacaaaaaaa	aaaaaaaaaa	2220
actcgagggg	gggcccggta	cccaattcgc	cctatagtga	gtcgtattac	a	2271

&lt;210&gt; 53

&lt;211&gt; 2769

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 53

ctgccaccgc	cgtccccccc	ctcccgtctc	cctcggggcg	ggctggggtcg	agctgcgatg	60
ccctcggact	tcattctcatt	gctcagcgcg	gacctagacc	tggaaatcgcc	caagtccctc	120
tactcgcgag	attctctgaa	gttacaccca	tcacagaatt	ttcatagagc	tggactattg	180
gaaggtcagc	aaagtaccat	tttaaagcat	attgtgttag	tatgcaaagg	ctttctctga	240
aatggagttt	aaatgggaaa	gtatgaaact	ttgcaaactc	acacaatgta	gttttctcct	300
aaagagtctg	atccttcttt	tagagcagct	gaatgtttca	atgggttttg	ttgctgcggt	360
tgatgtgctt	gttggtgtgt	ctatctgctt	tgagaaacat	tgaaaaaaaaa	atcattttta	420
attatgtgtt	gagtgttcta	ctttagaatt	ccagaccttc	atttttagctt	tttatcacaca	480
aaatagaatg	tgaaaaggcc	caaacattaa	catgaatata	aaagagtgc	accaaacaca	540
taattagttt	gaaaaatcag	aatttagaac	atttatatgc	tgagcatttc	aaaatatggt	600
aattatttat	tgcaaaagca	ttgatgccat	tgttattgca	ttatacttaa	gggtaaggaa	660
agccaagtaa	tagcgatgat	gattgagcta	tgcttagaag	aaaggggttt	agaaagtctt	720
ggaaatgtta	atagagtctg	tagattagtt	taaacagctt	acacttactt	tagaaaactt	780
ttggcctttg	gtgtgtgtgt	gttttaacat	tttaagacat	aaaatagtct	tgtgtttttg	840
tagattttga	agtgatttca	catatatcat	ctttaaatct	ttataatttt	ggaattggac	900
agcaaaggta	ctaatttccc	cattgtagag	aagagaaact	aaggctgaaa	ttaagaaacc	960
aaaaaaaaaa	aaaagtccca	cagtagaaga	aataataact	tctgcgtttc	agttagtgtc	1020
ctttatgcaa	agcggccacc	gtgaaataca	catgtaacca	gtaacaaaca	ttttatgggtc	1080
tactttaaaa	aggccctaaa	tatttaagtt	tgacttgaga	agattttttg	ctcccttgga	1140
ggtagtccgg	gatagtgggg	ggcaagtggg	ttttgtagtt	gagcttaaat	ttggatacca	1200
acttagctcc	ttattagttc	tgtgacttcg	ggcaagtgat	tttacctctc	taaactttta	1260
tattcttctc	gttaaaatgt	agctgtctct	taaatttagt	gagctaattg	gtattaaatt	1320
gcctagcata	gcgtctagct	taataaatgt	tagttttttc	cttttagtatc	agctaaattt	1380
ttttaacagt	ataatttaga	gcataataaa	tgttctcaaa	tataataattg	gttaaatttg	1440
ttttgtacgt	gtaagaagcc	catagatata	cattagtatg	aaaatcaa	atttttaaaa	1500
ctacaatttg	aagtgagaaa	tttttagcag	tgtttcttag	tgaagcagat	ttctggtttt	1560
tattcatatt	caatttgtga	ccatgttgta	gagacatatt	tgtctgtaca	aacctatctc	1620
atttcctgga	caacaaacca	tagagtgtca	taatcataga	gatgatcagg	gaggagaaag	1680
acatgaagaa	aaaaatgagt	gttaattatt	taagcaaata	tttttagtga	gctgatggaa	1740
atacagacaa	ccatttagca	tgttttattta	catgtttatg	cctaaatact	ttcggttaatt	1800
tttcatttta	gttaagggga	ttctcattta	aagtaaggct	acccttttag	aaattatata	1860
taaaatataa	catgtataag	tatcaattct	gtcatgccct	ttaaaatggt	ttgagacagg	1920
gtgtctttaa	aaagttgttt	gcttctgtaa	tgaagtgtct	aatctgtgcc	acctcattgc	1980
tatttctttt	ttattattga	cattttaaag	aatacataaa	agtagagaga	ataaaaataag	2040
ccccatatac	ccatcaccac	agtttaacca	gtatctaata	ttgcttcttt	tatcccttct	2100
tctctctcaa	gctggattat	tttaaaacaa	atatcaaaca	tcagatcact	tcatacataa	2160
gtatttctgt	attacagtga	ctttcttata	tttaattattg	atataattta	attgttaata	2220
taattattgt	attaattttt	aattactctg	tatagtttat	aaactatttc	ttggaatatt	2280
ctagtttatt	tcagcaaagt	ctcttgtgac	tctcctgccc	tcttttcaag	acagagtgc	2340
gcaggaggct	accgcttcta	gtatatagta	tatcctataa	cctatgttga	aattgtaggc	2400
catgctgcta	gattggagtt	catattgttg	ctctgacatt	tacttgtatg	gccttagcct	2460
aataatttaa	tcttttttgc	ttcaattttc	tcatctttta	acgggagcaa	taatagtgcc	2520
tcatacctagg	tttttgtgaa	gattatttga	ggcaaattcg	tgtaaaagcat	tttaatttagc	2580

38

acatgctaag	tgctcagtaa	atgttagcta	aagttaaaaa	atttttaaaa	cctaaatgct	2640
agctgtagag	tgtttactgt	gtgccaagaa	ctttacaaat	aagaactcct	gagggtgggca	2700
tatcacttgc	ggtcaggagt	ttgagaccag	cctggcttac	atggtgaaac	cccattctcta	2760
ctaaaaaaaa						2769

&lt;210&gt; 54

&lt;211&gt; 1389

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 54

ctgcaggatt	cggcagcagg	atctgttgac	cacagatgta	agaatttatt	tctggagcct	60
caattctaata	ccattgctct	acatgtgcac	tggtttcctg	gggctgccac	agcagaatac	120
cacagacttg	ggggcttaaa	caacaaagtg	attcctcaca	gttctggagg	ctccaagtct	180
gagatcaagg	agtcggcatg	tttggttcct	cctgaggcct	ctctccttgg	cctgcagatg	240
accacgttct	cactgtgttc	tcagctggcc	cttctctgtg	catgtacatc	cctgggtgtct	300
cttctcctct	ttgttgacta	taaggatacc	agtctgttg	gaccagagcc	ccactgtaaa	360
ggcctcatct	taacttaatc	cacttaaata	acttatttcc	acacacagta	agtctgaggt	420
tgagtctttt	gaattctggg	ggagacactg	cagcccatca	cgtatgttga	cgtgtgccac	480
cgccacactg	cctgggtact	gcagagtgtt	agtaagtttt	gaaattagaa	agtataagtg	540
ctccagttct	gttatttttc	aggattatta	gggctattct	gggtcccttc	catctccata	600
tgaatttgag	aatcagcttg	tttattcctg	ctaaaattta	gttgggattt	tttttttttt	660
gagatggagc	ctggctgtgt	caccaggctg	gtgtgcagtg	gtgcaatctc	gcctcactgc	720
aacctccacc	tcccgggttc	aagtgattct	cctgcctcag	cctcctgagt	agctgggact	780
acatgcgcac	gccaccacac	tcagctaatt	tttgtatttt	tagtagagac	gggggtttcac	840
catgttggcc	aggatgggtct	tgatctcttg	acctcgtgat	ccgcccacct	tggcctccca	900
aagtgtctggg	attacaggca	tgagccaccg	cgcctggcct	cagttgggat	tttcacaggg	960
ctcgttttga	gtctgcagat	caatttgggt	agtactgaca	tcttaaaatt	aagtctaccg	1020
atccacataa	gtgggatgct	agttaatatg	gacaagtga	tttattactc	ctccaaaaaa	1080
ctattttcac	aaactttata	tcaacatgag	aaatcatggt	gtatctttct	acctggaggg	1140
agcaagaaag	tcaggtgggtg	tgccttttgt	acatgggcaa	cccaatttct	tatctagact	1200
ttcttttgct	cgcctggaaa	aaaaaaaaaa	tgagggtattg	caaaaatgtg	ataattgctt	1260
tccttagtat	gaaaaactgt	taaaatgagt	aaagcctgtg	ctcaactaga	attctcacat	1320
ttccgttttg	atgtatccgc	atttcacctc	gtgccgaatt	cgatatcaag	cttatcgata	1380
ccgtcgacc						1389

&lt;210&gt; 55

&lt;211&gt; 748

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (15)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (17)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (29)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

<221> SITE  
 <222> (32)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (646)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (741)  
 <223> n equals a,t,g, or c

<400> 55  
 ggtgcggccg tttnantag ggatccccng gntgcaggaa ttcggcacga gcttattgka 60  
 tgtgctmcag tgacaggcat ctgggacttt tccaatctga agctgttagg aacaccacca 120  
 ccatgaatat tcttgtatgt gttttctggc tctgggggtgg agttgctggg tcatggggca 180  
 ggcacatttt catttttact tcagtaaaaa atgtcyagkg ggccagccac tgtgcctggc 240  
 ccaratgact tcttcacaag gagacacata ctgtgtgtat cagtcaggat ccaaccagga 300  
 gacaaaaccac acagtaattt aaacagcgat tgtttaatat acagaattgt taactatgat 360  
 aggggatttg agtaagagga actgggttact aagaaataaa gagaatgcta acgaatgtag 420  
 aaatagactg ggcacaatgg ctcacacctg taatcccagc actttgggag gcaaggcagg 480  
 tggatcacga ggtcaagaga tcgagaccat cctggccaac atggtgaaac cccgtcccta 540  
 ctacaaaaag tagctgggcg tgggtggcgtg tgctgttagt cccagctact caaggaggct 600  
 gaggcaggag aatagtttga acccaggagg cagagggtgc ggtganccga gatcttgcca 660  
 ctgcactcca gcctgggtga cagagagaga ctccatctca aaaaaaaaaa aaaaaaaaaa 720  
 actcgagggg gggcccggta nccaattc 748

<210> 56  
 <211> 4202  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (57)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (58)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (61)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (4137)  
 <223> n equals a,t,g, or c

<400> 56  
 accagctcaa agattcatga gtttcatcga gtcactgtga gtggagccca tgctggmntg 60  
 ntgccctctg tgtctgtgca tgcgcgtgtg tgtgtggcg tgtgtgcatt gctgggccag 120  
 cttgaaggga aggcccgta tgtccctgca ctctgttttg caagatgcca aacccagtt 180

ctgatggggc	tccaacagcc	aggctgtggt	cctttgacgt	tcctcacctg	ttgccaacct	240
atcccgtagt	gaactgaaac	cccaatgaag	acagaactgt	gcctggggag	atgcaatgag	300
gtgagggctg	aactcatcct	tttatatttc	ttttcaagat	tggatcagag	ctcatctcca	360
tccagtcttg	tttctatgaa	ggcttcaatc	tgtttccatg	caaatttgct	aatcagagcc	420
cagagctgct	gggtccctca	tctccctcat	ctattataga	ttgacttaca	gcaggagag	480
aatctcttta	gctcattcct	aatgggggtg	ggatcacaaat	atgggtctggt	ccaatctgca	540
tcttgttgtg	tcccaagacc	ctatctcctc	cccaacattc	ttattgcctt	tggtctccag	600
taaggaaacga	attggggggc	agggaggaga	acagggggga	tcaagaaggg	aaacccaatt	660
ccccctttga	aagtgggttc	tttgaactat	gtgtttgggg	gaagttcctc	tggatactaa	720
tttgaattta	tatacctcat	gttttggggg	tttgacgtat	atatatatat	atatatatat	780
atgcataatat	atttcataat	atttgggaagg	tttttggatgc	tagaaaaatg	gaaacaagag	840
aaccttcaaa	aatgggtactt	agatgggaac	tggaggccaa	tctttcataa	agccagcccc	900
atagctgctt	gctgttaggc	ctccagccat	tttgacattg	gggtggatag	tcgattcacc	960
tgctgtcgag	tcgattcacc	tgctgtcac	ccagtctctg	ggatgtgctg	gtgctgagcc	1020
tttgcctctt	ttccaaatgg	ttacagggat	gttgatcagc	tccaccagag	ggagctctga	1080
tgggaggaat	tgctctgcca	tccttgtccc	tgtgtctcct	gtcggcaggc	agccattgta	1140
tctcaccagc	agaccaggag	actggtccca	aggttactgc	accacagggc	aatttctctgc	1200
catagttagg	aaggaaacac	ctgaactaaa	tgggaagagac	atccctgcgg	tgtttaatat	1260
cacacccatg	ccctttgtca	ggttaccatg	tacagagatt	acttggagag	cctcatgccg	1320
tctctacctt	cgcacactgg	tcaagtatct	gctgagcttc	ttggccgcaa	ggatgcagaa	1380
ataggctgag	ggtccatggg	aagaaagaca	caatgaggca	gtaggagggtg	gggaagaaaa	1440
gaagacagac	tttcaaaatg	gaattaggca	ctggggagag	atcagtttcc	ccacatcagg	1500
gagaagaagg	tatagggtggg	gaaggggggtg	gccaggagca	gaaggaagaa	gactcaagat	1560
ggaaagggag	ccgctgtgcc	tgtggcaata	ccacttggag	aggctgactt	cataccttca	1620
agccttttcc	cctgggcttt	tgattgtgtc	tgtgccccct	ttcttgtcct	ctctgcagat	1680
gccagtagg	ggctacctca	tcctcgtgct	gttcttgtgt	ggctttcttg	gcagtaggga	1740
tcttgaattt	cctttctaac	actgtgcccg	gcaaggcggg	gagcattcct	ctgccctttg	1800
tcttgtgcca	acctggaaag	gtgcagtcta	gatttcagtg	agaaccctgc	cagctgagcc	1860
ctgtgcatct	actaccttga	cacagagtgt	tttcccacta	gaagctctgc	tctgctctcc	1920
tggcccaagt	aggggattcc	atgccttccc	tttcatggtc	ttagcaccag	cagcctagtt	1980
tctcccttcc	agagtctcca	gggatgacaa	attggattgg	agacaaacct	cgtcagatgc	2040
tcacccccta	aaagggtaat	tgtgtatttg	tggctgcgtg	tgctttgtg	ttttcattct	2100
cttcccattt	ttgtacattt	tggctttctc	tgtgggttta	tacttgggtca	aaagtactcg	2160
tcttgggtatt	gcactgttgt	gtgcatgaga	aaactggggg	aaggctcact	ggtacaagaa	2220
aggacccctg	acccctttcc	ttctctgtgg	tcccgccat	tagattgggg	gttctgggag	2280
aggcaggtga	atgtcctaag	tgaattgttc	tgtttgtaac	tggaaatgtt	ttgaagtctt	2340
tgggtgttgc	ccgtgaaagg	acatcgccac	ctggtgtcta	tgaggtgtct	ttgcagaaca	2400
ataaatggca	aatgaacaac	camaaaattg	ttacycttgt	tggccttctg	ctgtttgtag	2460
attagtgcac	ctatctgtga	gggatttggg	ttacctccct	gagtctgtaa	gcaaccacaa	2520
gccctgccac	tgggtggggg	aagtcctccc	ccaaccactt	aaaaacaaat	tttcccacat	2580
attaccccac	cccacacatt	tgacctggc	tagactttgt	ttgcctaaag	gaacagacca	2640
cattgtctgg	aaaatgagta	agtgaacgtg	tgggagaaaa	acacttttag	aatcacgaat	2700
attcactttt	aaaggctctt	ttgcctggct	gcaatatagt	gtgtgtttta	attatttaca	2760
ggctgttgtt	tctcaaataa	atgtttaata	ttaatcattc	ccaaactgac	aagaacacaa	2820
aaataaaaatg	caaatacaga	gccagctttg	tcacccaaat	ctgtgtctat	ttctgatagt	2880
ccatggaatg	tgggttttctt	ggaagccagg	gttggctctcc	ccacagaccc	caggctaagg	2940
tcaccagtta	ggaacccagg	acttggaaag	cagagctgtg	agctcttcca	tcagggatct	3000
gactccgcaa	aacgacttga	tgaatgcaat	tggcaaacctc	ccatgttcgg	acttcatatg	3060
catgagccgt	tggacagagg	gtttcttagt	atatacttta	atgcatgttt	atgtgcaatc	3120
ttgttagctg	gtatacaagt	ttgtgaagaa	cttctcattt	caataggcag	ttaatgtaat	3180
gcattaaaag	cctgggaatt	tggggctata	tttttccctt	ctgactcaat	aatcttcaaa	3240
gaattcatag	gaaagtcaat	acttgcagac	aagtgggttag	cttggctaaa	atgtacaaaa	3300
caccagaaac	ccacaaaaca	ctcagagggt	taggagaatg	ttttaatgct	taagaggcag	3360
gatcaagtga	agagggtaca	gaaatcagtg	tctctggctg	ggcagtcagg	agagcgggct	3420
caaattctgt	gactcacttc	tctgtgtctc	ggttggaaat	gaatgggtat	cctgggtccc	3480
accttcccac	acgctgtgat	acttcaaact	ccttgggtga	agggcctctt	ctcagcccaa	3540
gatcttgatt	gtgaacatta	acaaagagaa	cagtcaccc	ccacagaaga	taactcatta	3600
atgacatttg	attcagtgaa	taaaatatc	atttataaaa	atattgtagg	gggatcatga	3660
aagtagtgga	ggtaattaca	atcaggagag	attgggtatta	aaattgagca	aagtcaccaac	3720

41

tctcaccaga	tgacaattat	gcatectgct	agatgcccca	gggctgtcag	cctggaactg	3780
aaataaatgt	gttataagt	gtgctggatg	cctttttcag	ttcatttgaa	aacatggatt	3840
tgatcatgtc	agctcccttt	ctgctggaaa	aaaaagtagt	ttgcataatt	ggtgttaact	3900
actctgtttt	gattctacag	agtaagtaat	actcaaatgt	ggtcttactt	taacttcttg	3960
cctttgttac	ccccagaacc	atgcagacat	tgaaatgtgg	tgtgcgtgtg	tgtgtgtgtg	4020
tgtgtgtgtg	tgtgtgttca	tataacagga	ggacaggaaa	ggtaaggacc	cagaacaatg	4080
aagacttatt	gaaatgtggg	gggggtgtgt	gtgtgtgtgt	gttcatagaa	caggagnaca	4140
ggaaaggtaa	tgaccatcca	tggaagatga	aggggtagta	cttagccgga	cgcggtgggtc	4200
ga						4202

&lt;210&gt; 57

&lt;211&gt; 854

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (831)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (839)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (844)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (851)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 57

cgaaattact	cctcactaaa	gttgaacaaa	agctggagct	cgcgcgccctg	caggctcgaca	60
ctagtggatc	caaagaattc	ggcacgagcg	ggatcccggc	aaaatgcaga	ttccccaaaa	120
tttttgtaaa	tacagatgac	acttatgaag	agctccattt	aatcgtttat	aaggtaacaa	180
ctgttttctt	tccagcggtta	atgattgtgc	tgaagtggat	ctttcttgca	tgtgtacacg	240
agtgcattgt	caaacctctt	aaatgtttct	tggaaaagat	attggaagtt	ctgattatgg	300
taaaactcaa	aatgggtggt	cttccagcgt	aataagttta	ttttcagctc	cttttaaaca	360
gttctgttat	tagtgaaaga	ggaactgttt	aagattgtga	tttataaacg	tgtgaagtct	420
cacgtgctct	caaaccaaag	gctgtcagag	gttggtgctg	cctgttctcg	aaatggctct	480
ggatgggggc	cgtagccacg	tgtctgtgca	tatgctgctt	ttgctctgat	tttaaagctg	540
taggcttgct	aattccataa	ggatcgat	ttgkttctgt	caggacatgg	tcttgtaagg	600
atatgtactc	aggttgtgtt	tctaattaaa	ggcagttttt	gattcaagaa	agaagacgga	660
gcatgtgcac	gtgtttctcc	tctttctctc	ctgaggctgt	ggagaagtgt	tcattttataa	720
aggctcagaa	atgatgccgt	ggggggamcag	gaaggagcgg	agaactagtc	tcgagagtac	780
ttctagagcg	gccgcgggcc	catcgatttt	ccacccgggt	ggggtaccag	ntaagtgtng	840
aagnattccc	ntta					854

&lt;210&gt; 58

&lt;211&gt; 1455

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



42

<220>  
 <221> SITE  
 <222> (5)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (39)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (40)  
 <223> n equals a,t,g, or c

<400> 58  
 ggcanagctg ggatgggggtc agcaccacaga agccagccnn ctctgacagc ttcctcttttg 60  
 gccaagccct gcctctgtac agcctcgagt ggacagccag agggcacgag ggagcccaga 120  
 gcccaagatg gagccccagc tggggcctga ggctgccgcc ctccgccctg gctggctggc 180  
 cctgctgctg tgggtctcag ccctgagctg ttctttctcc ttgccagctt cttccctttc 240  
 ttctctgggtg ccccaagtca gaaccagcta caattttgga aggactttcc tcggtcttga 300  
 taaatgcaat gcctgcatcg ggacatctat ttgcaagaag ttctttaaag aagaaataag 360  
 atctgacaac tggctggctt ccacacctgg actgcctccc gattccttgc ttctttatcc 420  
 tgcaaatcac tcagatgatt ccaaaatctg gcgcctgtg gagatcttta gactggctcag 480  
 caaatatcaa aacgagatct cagacaggaa aatctgtgcc tctgcatcag ccccaaagac 540  
 ctgcagcatt gagcgtgtcc tgcggaaaac agagagggtc cagaaatggc tgcaggccaa 600  
 gcgcctcacg cgggacctgg tgcaggactg tcaccagggc cagagagaac taaagttcct 660  
 gtgtatgctg agataacacc agtgaaaaag cctggcatgg agcccagcac tgagaacttc 720  
 cagaaagtgt tagccttctc ccaactgtgt tataccaacc acattttcaa atagtaatca 780  
 ttaaagaggc ttctgcatca aaccttcaca tgcagctccc atgccacctc cagaattcac 840  
 caacacacag gccaccagc aacaggtacc ttgacacat attttttgat gacaatccaa 900  
 agccccggct ctttccacc acactgtggt cccctagatg gggctgttgc tgagcccacc 960  
 ccaatcccag atgtgatccc ccctgtgatc tacttctgg caagattcct ccagtcctgg 1020  
 acaggctctc cctatgagat agaacctgat aaggagctag ggcaattctg acaacattac 1080  
 caaaggccca cataacttct aaattttggt ctggtctgaa ggaaaacctg ttcttgccct 1140  
 agtgatggat gaactctctt atctctggct tctagaggga aaaaaagca tacctctttt 1200  
 actttttaag tacctccatc agagtcatga aatcacctgt caagactatc tatcttttat 1260  
 gtttccattc tggttaagaac tctttaaatg aggacactgc tgattgtctg tgatgttttt 1320  
 tgagcaaaca ctcgggggtg tggatgaaag ccaatcgag gtcaaagac tccttgggga 1380  
 agctacttct cctctattca gatttacta aaatcttcca agatgaaagc aaaaaaaaaa 1440  
 aaaaaaaaaa aaaaaa 1455

<210> 59  
 <211> 593  
 <212> DNA  
 <213> Homo sapiens

<400> 59  
 ccacgcttcc ggagctctta cttctccagc aacgctcttc agtacataat aagcttaact 60  
 gataaacaga atatttagaa aggtgagact tgggcttacc attgggttta aatcataggg 120  
 acctaggcg agggttcagg gcttctctgg agcagatatt gtcaagttca tggccttagg 180  
 tagcatgtat ctggtcttaa ctctgattgt agcaaaagt ctgagaggag ctgagccctg 240  
 ttgtggccca ttaaagaaca gggctcctcag gccctgcccg cttcctgtcc actgccccct 300  
 ccccatcccc agcccagccg agggaatccc gtgggttgc taccaccta taagggtggtt 360  
 tataagctgc tgtcctggcc actgcattca aattccaatg tgtacttcat agtgtaaaaa 420  
 tttatattat tgtgaggttt ttctgtcttt tttttttttt tttttttttg gtatattgct 480  
 gtatctactt taacttccag aaataaacgt tatataggaa ccgacaaaaa aaaaaaaaaa 540  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 593

<210> 60  
 <211> 496  
 <212> DNA  
 <213> Homo sapiens

<400> 60  
 aattcggcctt tgcagcggcc gcccgggcag gtcctgcttt ttgcttataa ttgacaacat 60  
 gtgcaaaaat accaaatttg tgcctctgtgc agtatgaaga attcagtgaa tattcattaa 120  
 tgtattagct tgttttgctc tctgttcata tatggctcta ttcttagaaa tataatttga 180  
 atgtgatctt tcaatagtct gaatatatta caaattatag ctatgtcttg tgaaaataac 240  
 ctcaaaaaga aaaatacgac tctgttgtct tacttgatat ttcttgccct agtaatgtac 300  
 ttgacattta tgttcctaag cagtgttaagt accagtagaa tttctctgtc aaactcaatg 360  
 atcatttagt acttttgtct tctcccatgt gcttgaagga aaaataaagt gtcactaccg 420  
 tatttcttgt tttcatcaaa aaataaaaat aatttaaaaa aaaaaaaaaa aaaaaaaaaa 480  
 aarrggsggg cccccc 496

<210> 61  
 <211> 1292  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (71)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (697)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1280)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1287)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (1291)  
 <223> n equals a,t,g, or c

<400> 61  
 aaacctcttc tataggtaaa gctggtagcg ctgcaggtag cggtccggaa ttcccggtc 60  
 gacccacgcg nccggaaga ggaacatag aggtgccaaa ggaacaaaga cataatgatg 120  
 tcatccaagc caacaagcca tgctgaagta aatgaaacca tacccaaccc ttaccaccca 180  
 agcagcttta tggctcctgg atttcaacag cctctgggtt caatcaactt agaaaaccaa 240  
 gctcaggggtg ctacgcgtgc tcagccctac ggcatcacat ctccgggaat ctttgctagc 300  
 agtcaaccgg gtcaaggaaa tatacaaatg ataaatccaa gtgtgggaac agcagtaatg 360  
 aactttaaaag aagaagcaaa ggactaggg gtgatccaga tcatgggttg attgatgcac 420  
 attggttttg gaattgtttt gtgtttaata tccttctctt ttagagaagt attaggtttt 480  
 gcctctactg ctgktattgg tggataccca ttctgggggtg gcctttcttt tattatctct 540

44

ggctctctct	ctgtgtcagc	atccaaggag	ctttcccggt	gtctgggtgaa	aggcagcctg	600
ggaatgaaca	ttggtaggtc	tatcttggcc	ttcattggag	tgattctgct	gctgggtggat	660
atgtgcatca	atgggggtarc	tggccaagac	tactggncg	tgctttctgg	aaaaggcatt	720
tcagccacgc	tgatgatctt	ctccytcttg	gagttcttcg	tagcttgtgc	cacagcccat	780
tttgccaacc	aagcaaacac	cacaaccaat	atgtctgtcc	tggttattcc	aaatatgtat	840
gaaagcaacc	ctgkgacacc	agcgtcttct	tcagctcctc	ccagatgcaa	caactactca	900
gctaattgcc	ctaaaagaaa	aaggggtatc	agtctaactc	catggagaaa	aactacttgc	960
aaaaacttct	taagaagatg	tctttttattg	tctacaatga	tttctagtct	ttaaaaaactg	1020
tgtttgagat	ttgttttttag	gttggtcgct	aatgatggct	gtatctccct	tcactgtctc	1080
ttcctacatt	accactacta	catgctggca	aaggtgaagg	atcagaggac	tgaaaaatga	1140
ttctgcaact	ctcttaaagt	tagaaatggt	tctgttcata	ttactttttc	cttaataaaa	1200
tgtcattaga	aacaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaagggc	ggccgctcta	1260
gaggatccaa	gcttacgtan	gcgtgcntgc	na			1292

&lt;210&gt; 62

&lt;211&gt; 398

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 62

ccccctggc	ccccccatta	atgccatggt	ttgggaggag	cckgatcccc	tgaaccccat	60
atttaacctc	tactgcccms	ggaaatgcc	tacattatct	ttccctaatt	ggaagtataa	120
ttagagtgat	gttggtaggg	tagaaaaaga	gggagtcact	tgatgctttc	aggttaatca	180
gagctatggg	tgctacaggc	ttgtctttct	aagtgcata	ttcttatcta	attctcagat	240
cagggtttga	aagmtwtggg	gggtctttta	gattttaatc	cctactttct	ttatggtaca	300
aatatgtaca	aaagaaaaag	gtcttatatt	cttttacaca	aattttataa	ttaattttga	360
actccttctg	tttaaaaaaa	aaaaaaaaaa	aaaaaaaaa			398

&lt;210&gt; 63

&lt;211&gt; 1202

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (282)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (596)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (607)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1200)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 63

gaattcggca	cgagattaag	ttgtgcactt	taattgggtg	aattgtacat	gtragttata	60
tatctractg	tagttgtwat	taaaaaacaa	caggaggcca	tgtgggctgc	taggagtagc	120
aatgtctgty	cccagccagc	aggtagagac	cagggctgga	cagagkagta	tgggctgtgc	180

45

tgcagattat	ttgtggtacc	caactgttgc	ataaaacagg	gtgtgatctc	ttgcattgct	240
atgcatgagt	ggattcccag	taaattgtgc	caggctgcct	gnatgatgtg	tggcttgtgc	300
tttggatcgt	aatgcttacc	tatgctactt	aagttacata	ccctgtggcc	tttgtggcca	360
ggactgtggg	ctactacctg	kagtgattcg	ttaggggaaa	ggaccacacag	cctgtgcagg	420
aggaaaaaag	catctctgag	tacaggggtg	atgagctgga	tgagctgccg	ggcaagagcc	480
acgcacaccc	aggtggtgag	tcttaaggat	aaggtggaat	ttgccccata	gctgtcctgg	540
acagaaactg	cccagagaag	aatgaatgga	ggacataggg	ctctgtggtc	ccaccntttt	600
ttggganacc	tgtgactggt	cctgtttacca	tgtcaactta	gccccaaacc	catctctgat	660
tgacttggtt	gcttattttg	gcacattctt	gctccacaca	gccacataca	tactggctgc	720
tcctcsaagg	ccaggcagat	gcagcagctg	ttggggccasc	aaagargaar	gtcctggaar	780
gttctggcct	gaacgctgca	tctgttgtgt	gacagccaca	actgctcagg	cttccttgtc	840
tgtgggtgca	ctgtggggag	gagtgttatg	ataagaacat	tggctctcag	tcttccttgg	900
ggagaagttt	ggcctcacgt	gggatttggg	cgttgccctt	aggaaggctc	tctgcatgtc	960
tagttccagt	ttgtactggg	aagaattaaa	aaagtctgcc	agcttcttta	gtttgtcctg	1020
tcttttgtga	tgattctttc	tgagatcccc	tcctatcagc	tcaggagtgg	gattttctgg	1080
agaaggaaaag	tgtttttcct	gttcctcact	gctcaccttg	gggcattcag	gaacatgggc	1140
ctgatgaatt	tgcttgaagg	cagtctgtaa	tcccatcact	ttgggagcca	aagaggcggn	1200
ca						1202

&lt;210&gt; 64

&lt;211&gt; 1517

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 64

gattacgcca	actcgaat	ttt	aaccctcact	aaagggaa	ca	aaagctggag	ctccaccg	cg	60
gtggcgccg	ctctaga	act	agtggatccc	cgggctgca	ggaattcg	gc	acgagggagc		120
ccagagccca	agatggag	cc	ccagctgggg	cctgaggctg	ccgccctccg	ccctggctgg			180
ctggccctgc	tgctgtggg	ct	cagccctg	agctgttctt	tctccttggc	agcttcttcc			240
ctttcttctc	tggtgccca	ag	tcagaacc	agctacaatt	ttggaaggac	tttctctggg			300
cttgataaat	gcaatgcctg	cat	cgggaca	tctattttgca	agaagtctt	taaagaaaga			360
aataagatct	gacaactggc	tg	gcttccca	ccttgggact	gcctcccgat	tcccttgctt			420
tcttatcctg	gcaaattact	cag	gatgatt	ccaaaatctg	gcgccctgtg	gagatcttta			480
gactggtcag	caaatatcaa	aac	gagatct	cagacaggaa	aatctgtgcc	tctgcatcag			540
ccccaaagac	ctgcagcatt	gag	cgtgtcc	tgcggaac	agagagggtt	cagaaatggc			600
tgagggccaa	gcgcctcacg	cc	ggacctgg	tgaggactg	tcaccagggc	cagagagaac			660
taaaagtctc	gtgtatgctg	gag	ataaac	cagtgaac	gccttggcat	ggagccccag			720
cactagagaac	ttccagaaag	tg	ttagcctt	ctcccaactg	tggtatacca	accacatttt			780
caaatagtaa	tcattaaaga	gg	cttctgca	tcaaaccctt	acatgcagct	cccatgccac			840
cctccagaat	tcaccaaac	ac	aggccac	cagcaacagg	cttacctttt	gcacaatatt			900
ttttgatgac	aatccaaagc	cc	ggctctt	tcccaccaca	ctgtgggtccc	ctagatgggg			960
ctgttgctga	gccacccca	at	ccagatg	tgatccccct	gtgatctact	tctggggccaa			1020
gattctccag	tctggacagg	tct	cccccta	tgagatagaa	cctgataagg	agctagggca			1080
attctgacaa	cattaccaa	gg	ccacata	acttctaaat	tttgggtctgg	tctgaaggaa			1140
aacctgttct	tgccctagt	g	atggatgaac	tctcttatct	ctggcttcta	gagggaaaaa			1200
aaagcatacc	tcttttactt	tt	taagtacc	tccatcagag	tcatagaatc	acctgtcaag			1260
actatctatc	ttttatgttt	cc	attctggg	aagaactctt	taaatagagga	cactgctgat			1320
tgctgggtgat	gttttttgag	caa	acactcg	ggggtatgga	tgaaagccaa	tcgcagggtca			1380
aatgactcct	tggggaagct	act	tctcctc	tattcagatt	tcactaaaat	cttccaagat			1440
gaaagcaaaa	aaaaaaaaa	aaaa	aaaaaaa	actcgagggg	gggcccgtac	ccaattcgcc			1500
ctatagttag	tcgtatt								1517

&lt;210&gt; 65

&lt;211&gt; 526

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

46

<220>  
 <221> SITE  
 <222> (66)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (106)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (484)  
 <223> n equals a,t,g, or c

<400> 65  
 ctctgacagc ttctctcttg gccaaagccct gcctctgtac agcctcgagt ggacagccag 60  
 aggtcnagac tggagcccag agccaagat ggagcccag ctgggncctg aggctgccgc 120  
 cctccgccct ggctggctgg ccctgctgct gtgggtctca gccctgagct gttctttctc 180  
 cttgccagct tcttcccttt cttctctggt gcccgaagtc agaaccagct acaatttttg 240  
 aaggactttc ctcggtcttg ataaatgcaa tgcctgcac gccacatcta tttgcaagaa 300  
 gttctttaaa gaagaaataa gatctgacaa ctggctggct tcccaccttg ggactgcctc 360  
 ccgattccct ttgstttctt atccttgcaa attactccar atgattycca aaatctggsg 420  
 sccttgtgga ratcttttaa ctggtcagca awtwtcaaac gaaatctcca aacaggaaat 480  
 cttntgcctc ctgcatccac ccccaaagaa ctgacacatt gacggt 526

<210> 66  
 <211> 664  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (31)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (63)  
 <223> n equals a,t,g, or c

<400> 66  
 caggctctca atacggactc actcataggg naaagctggt acgcctgcag gtaccgggtcc 60  
 ggnaattccc gggtcgaccc acgcgtcgr gagctcttac ttctccagca acrtcttca 120  
 gtacataata agcttaactg ataaacagaa tatttagaaa ggtgagactt gggcttacca 180  
 ttgggtttaa atcatagggc cctagggcga gggttcaggg cttctctgga gcagatattg 240  
 tcaagttcat ggccttaggt agcatgtatc tggctttaac tctgattgta gcaaaagtgc 300  
 tgagaggagc tgagccctgt tgtggcccat taaagaacag ggtcctcagg ccctgcccgc 360  
 ttccctgtcca ctgcccctc cccatccca gccagccga gggaatcccg tgggttgctt 420  
 acctacctat aagggtggtt ataagctgct gtcttgccca ctgcattcaa attccaatgt 480  
 gtacttcata gtgtaaaaat ttatattatt gtgaggtttt ttgtcttttt tttttttttt 540  
 ttttttggtt tattgtgta tctactttaa cttccagaaa taaacgttat atrggaaaaa 600  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 660  
 aaaa 664

<210> 67  
 <211> 156

47

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (156)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 67

```

Met Arg Leu Trp Lys Ala Val Val Val Thr Leu Ala Phe Met Ser Val
 1             5             10             15

Asp Ile Cys Val Thr Thr Ala Ile Tyr Val Phe Ser His Leu Asp Arg
          20             25             30

Ser Leu Leu Glu Asp Ile Arg His Phe Asn Ile Phe Asp Ser Val Leu
      35             40             45

Asp Leu Trp Ala Ala Cys Leu Tyr Arg Ser Cys Leu Leu Leu Gly Ser
      50             55             60

His His Trp Cys Gly Gln Glu Gln Cys Ala Gly Ala Pro Ala Ala Ala
 65             70             75             80

Gly Leu Val Ala Gly His His Pro Arg Val Pro Leu Arg Gly His Leu
          85             90             95

Cys His Gly Glu Ala Ala Ala Leu Leu Arg Gly Ala Gln Ala His Pro
      100             105             110

Gly Pro Leu Val Leu Gly Pro Val Arg Val Asp Val His Phe Thr Arg
      115             120             125

Arg Ile Leu Pro Ala Leu Val Ala Ala Val His Arg Ala Ala Arg His
      130             135             140

Pro Gly Pro Gly Ala Arg Gly Gly His Arg Gly Xaa
145             150             155

```

&lt;210&gt; 68

&lt;211&gt; 70

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (70)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 68

```

Met Ala Ala Arg His Leu Pro Gly Phe His Thr Tyr Thr Asn Leu Leu
 1             5             10             15

Phe Leu Leu Leu Pro Ser Leu Leu Met Gly Tyr Ser Glu Ser Pro Pro
      20             25             30

Pro Ile Thr Asp Ser Trp Ala Pro Phe Ile Ser Leu Thr His His Val
      35             40             45

```

48

Leu Ser Gln Ser Gln Ser Pro Leu Ser Ser Asn Cys Trp Ile Cys Leu  
 50 55 60

Ser Thr His Thr Gln Xaa  
 65 70

&lt;210&gt; 69

&lt;211&gt; 502

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (502)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 69

Met Trp Lys Leu Trp Arg Ala Glu Glu Gly Ala Ala Ala Leu Gly Gly  
 1 5 10 15

Ala Leu Phe Leu Leu Leu Phe Ala Leu Gly Val Arg Gln Leu Leu Lys  
 20 25 30

Gln Arg Arg Pro Met Gly Phe Pro Pro Gly Pro Pro Gly Leu Pro Phe  
 35 40 45

Ile Gly Asn Ile Tyr Ser Leu Ala Ala Ser Ser Glu Leu Pro His Val  
 50 55 60

Tyr Met Arg Lys Gln Ser Gln Val Tyr Gly Glu Ile Phe Ser Leu Asp  
 65 70 75 80

Leu Gly Gly Ile Ser Thr Val Val Leu Asn Gly Tyr Asp Val Val Lys  
 85 90 95

Glu Cys Leu Val His Gln Ser Glu Ile Phe Ala Asp Arg Pro Cys Leu  
 100 105 110

Pro Leu Phe Met Lys Met Thr Lys Met Gly Gly Leu Leu Asn Ser Arg  
 115 120 125

Tyr Gly Arg Gly Trp Val Asp His Arg Arg Leu Ala Val Asn Ser Phe  
 130 135 140

Arg Tyr Phe Gly Tyr Gly Gln Lys Ser Phe Glu Ser Lys Ile Leu Glu  
 145 150 155 160

Glu Thr Lys Phe Phe Asn Asp Ala Ile Glu Thr Tyr Lys Gly Arg Pro  
 165 170 175

Phe Asp Phe Lys Gln Leu Ile Thr Asn Ala Val Ser Asn Ile Thr Asn  
 180 185 190

Leu Ile Ile Phe Gly Glu Arg Phe Thr Tyr Glu Asp Thr Asp Phe Gln  
 195 200 205

His Met Ile Glu Leu Phe Ser Glu Asn Val Glu Leu Ala Ala Ser Ala

[illegible]

<210>	70
<211>	189



50

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (85)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (104)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (164)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (189)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 70

Met	Arg	Pro	Ala	Phe	Ala	Leu	Cys	Leu	Leu	Trp	Gln	Ala	Leu	Trp	Pro
1				5					10					15	

Gly	Pro	Gly	Gly	Gly	Glu	His	Pro	Thr	Ala	Asp	Arg	Ala	Gly	Cys	Ser
			20					25					30		

Ala	Ser	Gly	Ala	Cys	Tyr	Ser	Leu	His	His	Ala	Thr	Met	Lys	Arg	Gln
		35					40					45			

Ala	Ala	Glu	Glu	Ala	Cys	Ile	Leu	Arg	Gly	Gly	Ala	Leu	Ser	Thr	Val
	50					55					60				

Arg	Ala	Gly	Ala	Glu	Leu	Arg	Ala	Val	Leu	Ala	Leu	Leu	Arg	Ala	Gly
65					70					75					80

Pro	Gly	Pro	Gly	Xaa	Gly	Ser	Lys	Asp	Leu	Leu	Phe	Trp	Val	Ala	Leu
				85					90						95

Glu	Arg	Arg	Arg	Ser	His	Cys	Xaa	Leu	Glu	Asn	Glu	Pro	Leu	Arg	Gly
			100					105						110	

Phe	Ser	Trp	Leu	Ser	Ser	Asp	Pro	Gly	Gly	Leu	Glu	Ser	Asp	Thr	Leu
		115					120					125			

Gln	Trp	Val	Glu	Glu	Pro	Gln	Arg	Ser	Cys	Thr	Ala	Arg	Arg	Trp	Val
	130					135						140			

Leu	Pro	Gly	His	Arg	Trp	Gly	Arg	Ala	Arg	Ser	Trp	Lys	Glu	Met	Arg
145					150					155				160	

Cys	His	Leu	Xaa	Ala	Asn	Ala	Thr	Cys	Ala	Ser	Thr	Ser	Leu	Arg	Ser
			165						170					175	

Cys	Val	Leu	Arg	Arg	Ala	Pro	Gly	Pro	Pro	Leu	Thr	Xaa
		180						185				

<210> 71  
 <211> 486  
 <212> PRT  
 <213> Homo sapiens

<400> 71

```

Met Gln Pro Ser Gly Leu Glu Gly Pro Gly Thr Phe Gly Arg Trp Pro
 1           5           10           15

Leu Leu Ser Leu Leu Leu Leu Leu Leu Leu Gln Pro Val Thr Cys
          20           25           30

Ala Tyr Thr Thr Pro Gly Pro Pro Arg Ala Leu Thr Thr Leu Gly Ala
          35           40           45

Pro Arg Ala His Thr Met Pro Gly Thr Tyr Ala Pro Ser Thr Thr Leu
 50           55           60

Ser Ser Pro Ser Thr Gln Gly Leu Gln Glu Gln Ala Arg Ala Leu Met
 65           70           75           80

Arg Asp Phe Pro Leu Val Asp Gly His Asn Asp Leu Pro Leu Val Leu
          85           90           95

Arg Gln Val Tyr Gln Lys Gly Leu Gln Asp Val Asn Leu Arg Asn Phe
          100          105          110

Ser Tyr Gly Gln Thr Ser Leu Asp Arg Leu Arg Asp Gly Leu Val Gly
          115          120          125

Ala Gln Phe Trp Ser Ala Tyr Val Pro Cys Gln Thr Gln Asp Arg Asp
          130          135          140

Ala Leu Arg Leu Thr Leu Glu Gln Ile Asp Leu Ile Arg Arg Met Cys
          145          150          155          160

Ala Ser Tyr Ser Glu Leu Glu Leu Val Thr Ser Ala Lys Ala Leu Asn
          165          170          175

Asp Thr Gln Lys Leu Ala Cys Leu Ile Gly Val Glu Gly Gly His Ser
          180          185          190

Leu Asp Asn Ser Leu Ser Ile Leu Arg Thr Phe Tyr Met Leu Gly Val
          195          200          205

Arg Tyr Leu Thr Leu Thr His Thr Cys Asn Thr Pro Trp Ala Glu Ser
          210          215          220

Ser Ala Lys Gly Val His Ser Phe Tyr Asn Asn Ile Ser Gly Leu Thr
          225          230          235          240

Asp Phe Gly Glu Lys Val Val Ala Glu Met Asn Arg Leu Gly Met Met
          245          250          255

Val Asp Leu Ser His Val Ser Asp Ala Val Ala Arg Arg Ala Leu Glu
          260          265          270

```

```

<210> 72
<211> 88
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (88)
<223> Xaa equals stop translation

<400> 72
Met Val Ala Ser Gly Trp Leu Leu Leu Ala Gln Ala Ser Phe Leu Pro
  1             5             10             15
Leu Ala Pro Pro Gly Ala Leu Gly Ala Gly Cys Trp Met Asp Gly Arg
      20             25             30

```

53

Pro Leu Ala Pro Pro Gly Ala Leu Gly Ala Gly Cys Trp Met Gly Gly  
 35 40 45

Arg Pro Leu Ala Pro Pro Gly Ala Leu Gly Ala Gly Cys Trp Met Gly  
 50 55 60

Gly Arg His Gly Ala Pro Leu Leu Gly Cys Leu Cys Pro Ser Gly Leu  
 65 70 75 80

Cys Ser Ser Tyr Val Cys Leu Xaa  
 85

<210> 73  
 <211> 299  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (167)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 73  
 Met Met Ser Ser Lys Pro Thr Ser His Ala Glu Val Asn Glu Thr Ile  
 1 5 10 15

Pro Asn Pro Tyr Pro Pro Ser Ser Phe Met Ala Pro Gly Phe Gln Gln  
 20 25 30

Pro Leu Gly Ser Ile Asn Leu Glu Asn Gln Ala Gln Gly Ala Gln Arg  
 35 40 45

Ala Gln Pro Tyr Gly Ile Thr Ser Pro Gly Ile Phe Ala Ser Ser Gln  
 50 55 60

Pro Gly Gln Gly Asn Ile Gln Met Ile Asn Pro Ser Val Gly Thr Ala  
 65 70 75 80

Val Met Asn Phe Lys Glu Glu Ala Lys Ala Leu Gly Val Ile Gln Ile  
 85 90 95

Met Val Gly Leu Met His Ile Gly Phe Gly Ile Val Leu Cys Leu Ile  
 100 105 110

Ser Phe Ser Phe Arg Glu Val Leu Gly Phe Ala Ser Thr Ala Val Ile  
 115 120 125

Gly Gly Tyr Pro Phe Trp Gly Gly Leu Ser Phe Ile Ile Ser Gly Ser  
 130 135 140

Leu Ser Val Ser Ala Ser Lys Glu Leu Ser Arg Cys Leu Val Lys Gly  
 145 150 155 160

Ser Leu Gly Met Asn Ile Xaa Ser Ser Ile Leu Ala Phe Ile Gly Val  
 165 170 175

Ile Leu Leu Leu Val Asp Met Cys Ile Asn Gly Val Ala Gly Gln Asp

54

180	185	190
Tyr Trp Ala Val Leu Ser Gly Lys Gly Ile Ser Ala Thr Leu Met Ile		
195	200	205
Phe Ser Leu Leu Glu Phe Phe Val Ala Cys Ala Thr Ala His Phe Ala		
210	215	220
Asn Gln Ala Asn Thr Thr Thr Asn Met Ser Val Leu Val Ile Pro Asn		
225	230	235
Met Tyr Glu Ser Asn Pro Val Thr Pro Ala Ser Ser Ser Ala Pro Pro		
245	250	255
Arg Cys Asn Asn Tyr Ser Ala Asn Ala Pro Lys Arg Lys Arg Gly Ile		
260	265	270
Ser Leu Ile Ser Trp Arg Lys Thr Thr Cys Lys Asn Phe Leu Arg Arg		
275	280	285
Cys Leu Leu Leu Ser Thr Met Ile Ser Ser Leu		
290	295	

<210> 74  
 <211> 48  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (48)  
 <223> Xaa equals stop translation

<400> 74

Met Ala Leu His Pro Gly Ser Ser His Leu Leu Val Ala Val Pro Val	
1                      5                      10                      15	
Ser Trp Phe Leu Phe Cys Ile Pro Gly Ile Ser Phe Ile Thr Leu Ser	
20                      25                      30	
Trp Ser Tyr Gln Glu Ser Pro Val Ser Phe Leu Ser Val Glu Gly Xaa	
35                      40                      45	

<210> 75  
 <211> 44  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (44)  
 <223> Xaa equals stop translation

<400> 75

55  
Met Tyr Ser Leu Phe Leu Thr Cys Ile Phe Pro Phe Thr Leu Cys His  
1 5 10 15

Lys Lys Ile Leu Met Val Ile His Asp Phe Thr Gly Pro Val His Val  
20 25 30

Phe Pro Glu Lys Thr Val Leu Glu Trp Asn Tyr Xaa  
35 40

<210> 76

<211> 140

<212> PRT

<213> Homo sapiens

<400> 76

Met Cys Ala Met Tyr Leu Met Ile Lys Ala Phe Leu Pro Lys Met Leu  
1 5 10 15

Ala Gln Lys Ser Gly Asn Ile Ile Asn Met Ser Ser Val Ala Ser Ser  
20 25 30

Val Lys Gly Val Val Asn Arg Cys Val Tyr Ser Thr Thr Lys Ala Ala  
35 40 45

Val Ile Gly Leu Thr Lys Ser Val Ala Ala Asp Phe Ile Gln Gln Gly  
50 55 60

Ile Arg Cys Asn Cys Val Cys Pro Gly Thr Val Asp Thr Pro Ser Leu  
65 70 75 80

Gln Glu Arg Ile Gln Ala Arg Gly Asn Pro Glu Glu Ala Arg Asn Asp  
85 90 95

Phe Leu Lys Arg Gln Lys Thr Gly Arg Phe Ala Thr Ala Glu Glu Ile  
100 105 110

Ala Met Leu Cys Val Tyr Leu Ala Ser Asp Glu Ser Ala Tyr Val Thr  
115 120 125

Gly Asn Pro Val Ile Ile Asp Gly Gly Trp Ser Leu  
130 135 140

<210> 77

<211> 153

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (153)

<223> Xaa equals stop translation

<400> 77

Met Leu Val Val Cys Leu Leu Leu Ala Thr Gly Phe Cys Leu Phe Arg  
1 5 10 15

Gly Leu Ile Ala Leu Asp Cys Pro Ser Glu Leu Cys Arg Leu Tyr Thr

20 25 56 30  
 Gln Phe Gln Glu Pro Tyr Leu Lys Asp Pro Ala Ala Tyr Pro Lys Ile  
 35 40 45  
 Gln Met Leu Ala Tyr Met Phe Tyr Ser Val Pro Tyr Phe Val Thr Ala  
 50 55 60  
 Leu Tyr Gly Leu Val Val Pro Gly Cys Ser Trp Met Pro Asp Ile Thr  
 65 70 75 80  
 Leu Ile His Ala Gly Gly Leu Ala Gln Ala Gln Phe Ser His Ile Gly  
 85 90 95  
 Ala Ser Leu His Ala Arg Thr Ala Tyr Val Tyr Arg Val Pro Glu Glu  
 100 105 110  
 Ala Lys Ile Leu Phe Leu Ala Leu Asn Ile Ala Tyr Gly Val Leu Pro  
 115 120 125  
 Gln Leu Leu Ala Tyr Arg Cys Ile Tyr Lys Pro Glu Phe Phe Ile Lys  
 130 135 140  
 Thr Lys Ala Glu Glu Lys Val Glu Xaa  
 145 150

<210> 78  
 <211> 180  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (48)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (180)  
 <223> Xaa equals stop translation

<400> 78  
 Met Ala Ala Ala Ser Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu Leu  
 1 5 10 15  
 Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ala  
 20 25 30  
 Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Xaa  
 35 40 45  
 Pro Val Ser Ser Ala Ile Pro Arg Arg Val Cys Trp Ser Leu Leu Ser  
 50 55 60  
 Pro Arg Pro Thr Arg Pro Pro Gly Pro Ala Pro Cys Pro Leu Pro Ser  
 65 70 75 80  
 Ala Gly Arg Gly Ala Ala Gly Leu Gly Pro Leu Ala Gln Gln Pro Val

85 57 95  
 90  
 Ser Pro Ala Pro Ala Ser Pro Met Ala Pro Cys Ser Pro Arg Gly Phe  
 100 105 110  
 Pro Pro Ala His Gly Val Glu Pro Glu Ile Leu Ala Thr Met Pro Val  
 115 120 125  
 Leu Thr Ser His Pro Pro Thr Pro Ser Pro Cys Ser Leu Gly Thr Cys  
 130 135 140  
 Arg Leu Leu Ser Ser Leu Cys Ala Phe Val Pro Gly Gly Leu Thr Leu  
 145 150 155 160  
 Leu Ser Leu Ala Gly Leu Gly Gly Pro Val Gln Ala Pro Ala Ala Pro  
 165 170 175  
 Pro Ser Leu Xaa  
 180

<210> 79  
 <211> 70  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (70)  
 <223> Xaa equals stop translation

<400> 79  
 Met Leu Met Gly Ser Ile Leu Tyr Val Leu Phe Cys Val Trp Leu Leu  
 1 5 10 15  
 Gln Cys Ile Phe Glu Ile Tyr Pro His Cys Cys Val Tyr Pro Lys Cys  
 20 25 30  
 Val Leu Phe His Cys Gln Ile Met Phe Cys Tyr Met Asn Ile Leu Gln  
 35 40 45  
 Asn Ile Cys Leu Phe Ile Tyr Trp Trp Ile Phe Ala Phe Val Pro Val  
 50 55 60  
 Trp Gly Tyr Tyr Glu Xaa  
 65 70

<210> 80  
 <211> 191  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (191)  
 <223> Xaa equals stop translation

<400> 80



58  
 Met Arg Ala Cys Pro Trp Ala Gln Val Pro Leu Tyr Leu Leu Leu Asp  
 1 5 10 15  
 Gly His Leu Ala Val Ser Gln Ala Gly Val Met Ala Gly Val Ser Gly  
 20 25 30  
 Gly Arg Gly Gly Arg Arg Leu Arg Gly Pro Ile Thr Ser Arg Val Ile  
 35 40 45  
 Thr Ser Cys Gln Gln Pro Gly Val Gly Val Trp Val Ser Leu Arg Pro  
 50 55 60  
 Glu Leu Leu Asn Leu Glu Ser Leu Gly Val Ala Ala Lys Gly Val Tyr  
 65 70 75 80  
 Asp Lys His Val Ser Leu Asp Ile Ser Gly Glu Arg Ser Gly Ala Leu  
 85 90 95  
 Val Thr Phe Ser Lys Gly Cys Trp Ala Ser Glu Gln Ser Pro Pro Met  
 100 105 110  
 Ser Gln Pro Leu Gln Gly Pro Ser Leu Ser Leu His Pro Arg Pro Ser  
 115 120 125  
 Ala Ala Leu Val Met Ser Arg Arg Lys Val Leu Gly Cys Ala Gln Ser  
 130 135 140  
 Gln Glu Ser Lys Ile Cys Gln Ala Lys Ala Pro Gly Lys Ser Arg Arg  
 145 150 155 160  
 Ser Leu Gly Trp Pro Pro Gly Cys Gly Ala Ala Arg Ala Lys Thr Val  
 165 170 175  
 Asn Thr Ala Leu Gln Leu Ser Glu Pro Gln Phe Ser Asn Leu Xaa  
 180 185 190

&lt;210&gt; 81

&lt;211&gt; 166

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (127)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (166)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 81

Met Cys Leu Ser Leu Leu Ala Ala Leu Ala Cys Ser Ala Gly Asp Thr  
 1 5 10 15  
 Trp Ala Ser Glu Val Gly Pro Val Leu Ser Lys Ser Ser Pro Arg Leu  
 20 25 30

59

Ile Thr Thr Trp Glu Lys Val Pro Val Gly Thr Asn Gly Gly Val Thr  
           35                          40                          45

Val Val Gly Leu Val Ser Ser Leu Leu Gly Gly Thr Phe Val Gly Ile  
           50                          55                          60

Ala Tyr Phe Leu Thr Gln Leu Ile Phe Val Asn Asp Leu Asp Ile Ser  
           65                          70                          75                          80

Ala Pro Gln Trp Pro Ile Ile Ala Phe Gly Gly Leu Ala Gly Leu Leu  
                           85                          90                          95

Gly Ser Ile Val Asp Ser Tyr Leu Gly Ala Thr Met Gln Tyr Thr Gly  
                           100                          105                          110

Leu Asp Glu Ser Thr Gly Met Val Val Asn Ser Pro Thr Asn Xaa Ala  
           115                          120                          125

Arg His Ile Ala Gly Lys Pro Ile Leu Asp Asn Asn Ala Val Asn Leu  
           130                          135                          140

Phe Ser Ser Val Leu Ile Ala Leu Leu Leu Pro Thr Ala Ala Trp Gly  
           145                          150                          155                          160

Phe Trp Pro Arg Gly Xaa  
                           165

<210> 82  
 <211> 42  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (42)  
 <223> Xaa equals stop translation

<400> 82  
 Met Cys Gly Leu Val Ile Leu Trp Pro Cys Ile Met Thr Leu Phe Ser  
       1                          5                          10                          15

Ser Leu Ser Thr Gly Asp Val Leu Leu Pro Cys Lys Ile Leu Val Gly  
           20                          25                          30

Leu Arg Val Phe Ile Gly Ala Arg Val Xaa  
           35                          40

<210> 83  
 <211> 49  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (49)  
 <223> Xaa equals stop translation

60

&lt;400&gt; 83

Met Cys Phe Pro Ala Cys Leu Cys Ser Pro Leu Thr Cys Leu Leu Ser  
1 5 10 15

Val Trp Lys Pro Gly Leu Ala His Ala Val Val His Cys Met Leu Glu  
20 25 30

Pro Val Glu Phe Ala Arg Val Val Gln Tyr Glu Ala Gly His Val Leu  
35 40 45

Xaa

&lt;210&gt; 84

&lt;211&gt; 57

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (57)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 84

Met Leu Ile Ala Lys Leu Pro Val Leu Glu Ser Ile Cys Phe Phe Met  
1 5 10 15

Leu Phe Leu Asn Pro Leu Val Ile Leu Leu Ser Leu Asn Asn Ala Leu  
20 25 30

Pro Leu Val Phe His Pro His Ser Glu Phe Leu Glu Asp His Asn Arg  
35 40 45

Gly Asp Thr Leu Pro Ser Ile Val Xaa  
50 55

&lt;210&gt; 85

&lt;211&gt; 43

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (43)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 85

Met Leu Val Ala Thr Ala Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala  
1 5 10 15

Ile Leu Ala Gln Leu Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu  
20 25 30

Thr Ile Trp Tyr Val Arg Phe Leu Trp Glu Xaa  
35 40

61

<210> 86  
<211> 41  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (41)  
<223> Xaa equals stop translation

<400> 86  
Met Leu Leu Leu Trp Ala Phe Ser Gly Val Cys Ala Val Pro Ala Arg  
1 5 10 15  
Ala Thr Pro Val Pro Ser Ser Phe Cys Pro Gln Gly Pro Ser Leu Cys  
20 25 30  
Pro Lys Gln Pro Ala Ser Leu Ala Xaa  
35 40

<210> 87  
<211> 74  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (74)  
<223> Xaa equals stop translation

<400> 87  
Met His Ala Tyr Ala Cys Val Cys Ala Cys Met Leu Val Cys Val Cys  
1 5 10 15  
Val Cys Val Cys Arg Ala Leu Val Ile Pro Thr Glu Gln Arg His Arg  
20 25 30  
Arg Val Ala His Gly Arg Thr Ser Asp Ser Thr Leu Pro Cys Thr Val  
35 40 45  
Lys Ile Trp Pro Ser Glu Arg Gly Asp Gly Arg Gly Glu Arg Gly Glu  
50 55 60  
Arg Arg Arg Gly Thr Asp Trp Arg Gly Xaa  
65 70

<210> 88  
<211> 47  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (34)  
<223> Xaa equals any of the naturally occurring L-amino acids  
  
<220>

62

&lt;221&gt; SITE

&lt;222&gt; (47)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 88

Met His His Pro Asn Leu Cys Leu His Phe His Ala Ala Phe Ser Leu  
 1 5 10 15

Cys Val His Gly Cys Leu Cys Val Gln Phe Phe Pro Phe Tyr Lys Asp  
 20 25 30

Thr Xaa His Ile Gly Leu Glu Pro Thr Leu Met Thr Ser Ser Xaa  
 35 40 45

&lt;210&gt; 89

&lt;211&gt; 63

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (63)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 89

Met Leu Phe Leu Asn Val Ile Leu Phe Ser Leu Thr Val Phe Thr Leu  
 1 5 10 15

Ile Ser Thr Ala His Thr Leu Asp Arg Ala Val Arg Ser Asp Trp Leu  
 20 25 30

Leu Leu Val Leu Ile Tyr Ala Cys Leu Glu Glu Leu Ile Pro Glu Leu  
 35 40 45

Ile Phe Asn Leu Tyr Cys Gln Gly Asn Ala Thr Leu Phe Phe Xaa  
 50 55 60

&lt;210&gt; 90

&lt;211&gt; 70

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (20)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (70)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 90

Met Leu Leu Lys Leu His Thr Leu Trp Pro Leu Trp Pro Gly Leu Trp  
 1 5 10 15

Ala Thr Thr Xaa Ser Asp Ser Leu Gly Glu Arg Thr His Ser Leu Cys

20 25 63 30  
 Arg Arg Lys Lys Ala Ser Leu Ser Thr Gly Trp Met Ser Trp Met Ser  
 35 40 45  
 Cys Arg Ala Arg Ala Thr His Thr Gln Val Val Ser Leu Lys Asp Lys  
 50 55 60  
 Val Glu Phe Ala Pro Xaa  
 65 70

<210> 91  
 <211> 57  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (57)  
 <223> Xaa equals stop translation

<400> 91  
 Met Lys Glu Ser Arg Lys Met Leu Trp Val Phe Lys Met Leu Phe Phe  
 1 5 10 15  
 Lys Ile Val Leu Trp Val Asn Leu Leu Ser Ala Ala Leu Ser Cys Ile  
 20 25 30  
 Gln Lys Gln Met Leu Gly Ile Ala Pro Gln Lys Cys Val Pro Lys Leu  
 35 40 45  
 Cys Phe Gln Leu Tyr Ile Met Arg Xaa  
 50 55

<210> 92  
 <211> 68  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (68)  
 <223> Xaa equals stop translation

<400> 92  
 Met Tyr Phe Leu Leu Ser Val Thr Ser Glu Ser Val Trp Arg Ser Trp  
 1 5 10 15  
 Thr Leu Thr Phe His Ser Phe Ala Ile Leu Ser Leu Arg Cys Trp Thr  
 20 25 30  
 Ser Leu Leu Leu Ile Pro Leu Thr Ser Cys Asn Phe Ser Ser Pro  
 35 40 45  
 Ser Trp Arg Met Thr Ala Ser Gln Val Pro Ser Lys Arg Lys Ala Ser  
 50 55 60

64

Met Thr Leu Xaa  
65

<210> 93  
<211> 45  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (45)  
<223> Xaa equals stop translation

<400> 93  
Met Lys Gly Trp Pro Val Phe Leu Leu Val Gln Ala Val Thr Phe Leu  
1 5 10 15  
Ser Val Ala Gln Ser Gly Ala Met Ala Cys Ala Ala Ser Gly Val Val  
20 25 30  
Tyr Ser Val Asp Val Pro Ala Cys Ser Ser Arg Ser Xaa  
35 40 45

<210> 94  
<211> 55  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (55)  
<223> Xaa equals stop translation

<400> 94  
Met Val Leu Ser Pro Trp Ala Cys Leu Phe Val Val Phe Phe Pro Tyr  
1 5 10 15  
Ile Gln Ser Ser Leu Arg Ser Asp Lys His Leu Gln Leu Ser Asn Ile  
20 25 30  
Leu Pro Thr Pro Ser His His Ile His Leu Pro Ala Ser Ile Cys Ile  
35 40 45  
Gln Leu Arg Ala Gly Asn Xaa  
50 55

<210> 95  
<211> 41  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (41)  
<223> Xaa equals stop translation

65

&lt;400&gt; 95

Met Cys Glu Tyr Val Leu Leu Leu Tyr Ile Val Leu Leu Cys Asn Arg  
 1 5 10 15

Ser Tyr Ala Val Phe Thr Gln Cys Val Leu Arg Ser Ser Pro Ile Asp  
 20 25 30

Ser Ser Arg Asn Ala Val Leu Leu Xaa  
 35 40

&lt;210&gt; 96

&lt;211&gt; 41

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (41)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 96

Met Thr Thr Pro Gly Leu Leu Ile Leu Phe Leu Ala His Val Cys Leu  
 1 5 10 15

Val Asn His Gln Gln Ala Ala Glu Pro Gly Trp Lys Gln His Cys Cys  
 20 25 30

Asn Trp Glu Gly His Arg Val Leu Xaa  
 35 40

&lt;210&gt; 97

&lt;211&gt; 50

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (14)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (50)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 97

Met Leu Cys His Val Tyr Leu Leu Leu Val Gly His Ala Xaa Phe Ser  
 1 5 10 15

Val Gly Leu Met Gly Gln Arg Lys Leu Arg Cys Ser Ile Asn Ser Ala  
 20 25 30

Leu Arg Ser Ala Val Ser Ser Ala Trp Asn Ser Ser Ile Cys Phe Asn  
 35 40 45

Ser Xaa  
 50



<210> 98  
<211> 58  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (58)  
<223> Xaa equals stop translation

<400> 98  
Met Ser Glu Trp Cys Gln Pro Asp Gln Ile Leu Leu Gln Phe Pro Val  
1 5 10 15  
Leu Ala Thr Met Ser Val Ala Phe Leu Ile Gln Arg Cys Phe Cys Phe  
20 25 30  
Trp Trp Phe Val Leu Asn Ala Phe Ser Ile Pro Ser Gly Thr Glu Lys  
35 40 45  
Lys Arg Ile Val Phe Lys Lys Trp Leu Xaa  
50 55

<210> 99  
<211> 52  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (52)  
<223> Xaa equals stop translation

<400> 99  
Met Lys Val Val Val Val Met Val Val Ile Leu Val Val Val Thr Leu  
1 5 10 15  
Val Val Val Val Met Val Val Ile Leu Val Met Val Val Met Val Val  
20 25 30  
Ala Leu Val Thr Leu Thr Trp Gly Pro Val Ala Val Thr Val Asp Ala  
35 40 45  
Gly Ser Trp Xaa  
50

<210> 100  
<211> 45  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (45)  
<223> Xaa equals stop translation

67

&lt;400&gt; 100

Met Pro His Phe Leu Arg Trp Leu Leu Thr Thr Phe Arg Ile Arg Ala  
1 5 10 15

Ser Cys Gly Ser Thr Pro Cys Trp Ser Pro Ser His Leu Gly Cys Leu  
20 25 30

Gln Pro Ala Leu Pro Arg Asp Leu Ser His Leu Glu Xaa  
35 40 45

&lt;210&gt; 101

&lt;211&gt; 58

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (58)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 101

Met Ser Thr Lys Ile Leu Gln Phe Leu Phe Ser Ser Cys Cys Trp Val  
1 5 10 15

Pro Pro Met Leu Phe Leu Phe Lys Asn Thr Lys Cys Arg Thr Ser Leu  
20 25 30

Leu Tyr Cys Phe Tyr Phe Ile Leu Leu Thr Cys Ser Leu Ser Glu Tyr  
35 40 45

Asp Ser Leu Leu Ser Ser Lys Val Phe Xaa  
50 55

&lt;210&gt; 102

&lt;211&gt; 41

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (41)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 102

Met Phe Trp Phe Trp Phe Leu Leu Ser Leu Ser Phe Gln Gln Val Glu  
1 5 10 15

Gln Gln Gln Val Phe Gln Cys Ile Cys Cys Thr Arg Thr Lys Tyr Lys  
20 25 30

Ser Val Trp His Gln Lys Ser Lys Xaa  
35 40

&lt;210&gt; 103

&lt;211&gt; 143

68

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (104)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (105)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (115)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (143)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 103

Met	Thr	Leu	Ile	Glu	Val	Leu	Val	Ser	Val	Leu	Ile	Leu	Ala	Val	Gly
1				5					10					15	

Leu	Leu	Arg	Ala	Ala	Val	Ile	Gln	Leu	Asn	Ala	Leu	Lys	Tyr	Thr	Asp
			20					25					30		

Ser	Ser	Arg	Met	Thr	Ser	Gln	Ala	Ser	Phe	Ile	Ala	Tyr	Asp	Met	Leu
		35					40						45		

Asp	Arg	Ile	Arg	Ala	Asn	Ser	Gly	Ala	Asp	Tyr	Ser	Trp	Gly	Gln	Gly
	50					55						60			

Glu	Arg	Ala	Pro	Ser	Thr	Thr	Ser	Val	Ala	Ser	Val	Arg	Asp	Leu	Asp
65					70					75				80	

Leu	His	Asp	Phe	Glu	Ala	Asn	Ile	Val	Gly	Phe	Ala	Gly	Glu	Ser	Ala
			85						90					95	

Lys	Gly	Ser	Val	Ala	Val	Asn	Xaa	Xaa	Glu	Val	Thr	Ile	Ser	Ile	Ser
			100					105						110	

Trp	Asp	Xaa	Ser	Arg	Gly	Ala	Asn	Ala	Gln	Gly	Thr	Arg	Glu	Thr	Phe
		115					120					125			

Thr	Leu	Thr	Ser	Arg	Val	Ala	Val	Asp	Pro	Arg	Val	Leu	Pro	Xaa	
	130					135						140			

&lt;210&gt; 104

&lt;211&gt; 44

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

69

&lt;222&gt; (44)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 104

Met Ala Phe Phe Phe Ala Leu Phe Val Ile Phe Phe Val Ile Val Val  
 1 5 10 15

Gln Met Glu Ser His Ser Gly Leu Gly Lys Lys Ser Lys Ile Leu Ser  
 20 25 30

Gly Gly Gln Gly Glu Glu Val Tyr Phe Leu Asp Xaa  
 35 40

&lt;210&gt; 105

&lt;211&gt; 63

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (63)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 105

Met Tyr Phe Tyr Leu Ala Val Lys Pro Pro Leu Leu Trp Ala Arg Pro  
 1 5 10 15

Gln Val Ser Cys Arg Leu Ser Val Ser Leu Ala Trp Ser Tyr His Leu  
 20 25 30

His Leu Trp Ala Leu Phe Leu Phe Ser Ile Leu Leu Gln Cys Arg Ala  
 35 40 45

Arg Phe Leu Leu Leu Val Leu Ser Gln Thr Gln Asp Leu Xaa  
 50 55 60

&lt;210&gt; 106

&lt;211&gt; 283

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (283)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 106

Met Gly Ser Pro Gly Met Val Leu Gly Leu Leu Val Gln Ile Trp Ala  
 1 5 10 15

Leu Gln Glu Ala Ser Ser Leu Ser Val Gln Gln Gly Pro Asn Leu Leu  
 20 25 30

Gln Val Arg Gln Gly Ser Gln Ala Thr Leu Val Cys Gln Val Asp Gln  
 35 40 45

Ala Thr Ala Trp Glu Arg Leu Arg Val Lys Trp Thr Lys Asp Gly Ala

50                      55                      70                      60  
 Ile Leu Cys Gln Pro Tyr Ile Thr Asn Gly Ser Leu Ser Leu Gly Val  
 65                      70                      75                      80  
 Cys Gly Pro Gln Gly Arg Leu Ser Trp Gln Ala Pro Ser His Leu Thr  
                     85                      90                      95  
 Leu Gln Leu Asp Pro Val Ser Leu Asn His Ser Gly Ala Tyr Val Cys  
                     100                      105                      110  
 Trp Ala Ala Val Glu Ile Pro Glu Leu Glu Glu Ala Glu Gly Asn Ile  
                     115                      120                      125  
 Thr Arg Leu Phe Val Asp Pro Asp Asp Pro Thr Gln Asn Arg Asn Arg  
 130                      135                      140  
 Ile Ala Ser Phe Pro Gly Phe Leu Phe Val Leu Leu Gly Val Gly Ser  
 145                      150                      155                      160  
 Met Gly Val Ala Ala Ile Val Trp Gly Ala Trp Phe Trp Gly Arg Arg  
                     165                      170                      175  
 Ser Cys Gln Gln Arg Asp Ser Gly Asn Ser Pro Gly Asn Ala Phe Tyr  
                     180                      185                      190  
 Ser Asn Val Leu Tyr Arg Pro Arg Gly Ala Pro Lys Lys Ser Glu Asp  
                     195                      200                      205  
 Cys Ser Gly Glu Gly Lys Asp Gln Arg Gly Gln Ser Ile Tyr Ser Thr  
 210                      215                      220  
 Ser Phe Pro Gln Pro Ala Pro Arg Gln Pro His Leu Ala Ser Arg Pro  
 225                      230                      235                      240  
 Cys Pro Ser Pro Arg Pro Cys Pro Ser Pro Arg Pro Gly His Pro Val  
                     245                      250                      255  
 Ser Met Val Arg Val Ser Pro Arg Pro Ser Pro Thr Gln Gln Pro Arg  
                     260                      265                      270  
 Pro Lys Gly Phe Pro Lys Val Gly Glu Glu Xaa  
                     275                      280

&lt;210&gt; 107

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (98)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 107

Met Cys Lys Leu Cys Phe Tyr Leu Tyr Leu Cys Thr Trp Phe Pro Phe  
 1                      5                      10                      15

71

Gly Ala Ser Gly Leu Phe Trp Asp Lys Trp Cys Leu Pro Arg His Leu  
                   20                  25                  30

Pro Val Val Ser Gly Gln Glu Gln Leu Ser Ser Ser Leu Pro Ala Ala  
                   35                  40                  45

Leu Leu Phe Leu Gly Arg Arg Trp Arg Pro Pro Leu Arg Val Ser Pro  
                   50                  55                  60

Gly Leu Ser Phe Arg Gly Gly Arg Ala Gly Glu Pro Gln Gly Trp Gly  
                   65                  70                  75                  80

Asp Ser Trp Glu Met Glu Val Ala Pro Ala Pro Leu Asp Gln Tyr Trp  
                   85                  90                  95

Leu Xaa

<210> 108  
 <211> 62  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (62)  
 <223> Xaa equals stop translation

<400> 108  
 Met Cys Leu Leu Leu Leu Trp Leu Thr Thr Phe Gln Arg Thr Ser Gly  
       1                  5                  10                  15

Ala Leu Arg Arg Gly Gly Leu Ser Ser Pro Ala Trp Ala Met Arg Ser  
                   20                  25                  30

Pro Ser Val Tyr Ser Thr Gln Thr Pro Ser Pro Met Met Ser Thr Gly  
                   35                  40                  45

Thr Leu Arg Gly Leu Ser Gly Ala Met Cys Asn Leu Ser Xaa  
                   50                  55                  60

<210> 109  
 <211> 47  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (47)  
 <223> Xaa equals stop translation

<400> 109  
 Met Lys Leu Cys Lys Leu Thr Gln Cys Ser Phe Leu Leu Lys Ser Leu  
       1                  5                  10                  15

Ile Leu Leu Leu Glu Gln Leu Asn Val Ser Met Gly Phe Val Ala Ala  
                   20                  25                  30

72

Phe Asp Val Leu Val Gly Cys Ser Ile Cys Phe Glu Lys His Xaa  
                   35                                  40                                  45

<210> 110  
 <211> 47  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (47)  
 <223> Xaa equals stop translation

<400> 110  
 Met Thr Thr Phe Ser Leu Cys Ser Gln Leu Ala Leu Leu Cys Ala Cys  
       1                                  5                                  10                                  15

Thr Ser Leu Val Ser Leu Pro Pro Phe Val Asp Tyr Lys Asp Thr Ser  
                   20                                  25                                  30

Pro Val Gly Pro Glu Pro His Cys Lys Gly Leu Ile Leu Thr Xaa  
                   35                                  40                                  45

<210> 111  
 <211> 42  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (32)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (33)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (41)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (42)  
 <223> Xaa equals stop translation

<400> 111  
 Met Asn Ile Leu Val Cys Val Phe Trp Leu Trp Gly Gly Val Ala Gly  
       1                                  5                                  10                                  15

Ser Trp Gly Arg His Ile Phe Ile Phe Thr Ser Val Lys Asn Val Xaa  
                   20                                  25                                  30

Xaa Ala Ser His Cys Ala Trp Pro Xaa Xaa

35 40 73

<210> 112  
<211> 41  
<212> PRT  
<213> Homo sapiens

<400> 112  
Met Gly Gly Ile Ala Leu Pro Ser Leu Ser Leu Cys Leu Leu Ser Ala  
1 5 10 15  
Gly Ser His Cys Ile Ser Pro Ala Asp Gln Glu Thr Gly Pro Lys Val  
20 25 30  
Thr Ala Pro Gln Gly Asn Phe Leu Pro  
35 40

<210> 113  
<211> 44  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (44)  
<223> Xaa equals stop translation

<400> 113  
Met Ile Val Leu Lys Trp Ile Phe Leu Ala Cys Val His Glu Cys Met  
1 5 10 15  
Cys Lys Pro Leu Lys Cys Phe Leu Glu Lys Ile Leu Glu Val Leu Ile  
20 25 30  
Met Val Lys Leu Lys Met Gly Val Leu Pro Ala Xaa  
35 40

<210> 114  
<211> 182  
<212> PRT  
<213> Homo sapiens

<400> 114  
Met Glu Pro Gln Leu Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp  
1 5 10 15  
Leu Ala Leu Leu Leu Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu  
20 25 30  
Pro Ala Ser Ser Leu Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr  
35 40 45  
Asn Phe Gly Arg Thr Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile  
50 55 60  
Gly Thr Ser Ile Cys Lys Lys Phe Phe Lys Glu Glu Ile Arg Ser Asp



```
<210> 115
<211> 81
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (81)
<223> Xaa equals stop translation
```

```

<400> 115
Met Ala Leu Gly Ser Met Tyr Leu Val Leu Thr Leu Ile Val Ala Lys
  1             5             10             15
Val Leu Arg Gly Ala Glu Pro Cys Cys Gly Pro Leu Lys Asn Arg Val
          20             25             30
Leu Arg Pro Cys Pro Leu Pro Val His Cys Pro Leu Pro Ile Pro Ser
          35             40             45
Pro Ala Glu Gly Ile Pro Trp Val Ala Tyr Leu Pro Ile Arg Trp Phe
          50             55             60
Ile Ser Cys Cys Pro Gly His Cys Ile Gln Ile Pro Met Cys Thr Ser
  65             70             75             80

```

```
<210> 116
<211> 49
<212> PRT
<213> Homo sapiens
```

75

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (49)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 116

Met	Ser	Cys	Glu	Asn	Asn	Leu	Lys	Lys	Lys	Asn	Thr	Thr	Leu	Leu	Ser
1				5					10					15	

Tyr	Leu	Ile	Phe	Leu	Ala	Leu	Val	Met	Tyr	Leu	Thr	Phe	Met	Phe	Leu
			20					25					30		

Ser	Ser	Val	Ser	Thr	Ser	Arg	Ile	Ser	Leu	Ser	Asn	Ser	Met	Ile	Ile
		35					40					45			

Xaa

&lt;210&gt; 117

&lt;211&gt; 204

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (31)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (93)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (99)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (115)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (151)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (204)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 117

Met	Val	Gly	Leu	Met	His	Ile	Gly	Phe	Gly	Ile	Val	Leu	Cys	Leu	Ile
1				5					10					15	

Ser	Phe	Ser	Phe	Arg	Glu	Val	Leu	Gly	Phe	Ala	Ser	Thr	Ala	Xaa	Ile
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

20 76 30  
 Gly Gly Tyr Pro Phe Trp Gly Gly Leu Ser Phe Ile Ile Ser Gly Ser  
 35 40 45  
 Leu Ser Val Ser Ala Ser Lys Glu Leu Ser Arg Cys Leu Val Lys Gly  
 50 55 60  
 Ser Leu Gly Met Asn Ile Gly Arg Ser Ile Leu Ala Phe Ile Gly Val  
 65 70 75 80  
 Ile Leu Leu Leu Val Asp Met Cys Ile Asn Gly Val Xaa Gly Gln Asp  
 85 90 95  
 Tyr Trp Xaa Val Leu Ser Gly Lys Gly Ile Ser Ala Thr Leu Met Ile  
 100 105 110  
 Phe Ser Xaa Leu Glu Phe Phe Val Ala Cys Ala Thr Ala His Phe Ala  
 115 120 125  
 Asn Gln Ala Asn Thr Thr Thr Asn Met Ser Val Leu Val Ile Pro Asn  
 130 135 140  
 Met Tyr Glu Ser Asn Pro Xaa Thr Pro Ala Ser Ser Ser Ala Pro Pro  
 145 150 155 160  
 Arg Cys Asn Asn Tyr Ser Ala Asn Ala Pro Lys Arg Lys Arg Gly Ile  
 165 170 175  
 Ser Leu Ile Ser Trp Arg Lys Thr Thr Cys Lys Asn Phe Leu Arg Arg  
 180 185 190  
 Cys Leu Leu Leu Ser Thr Met Ile Ser Ser Leu Xaa  
 195 200

&lt;210&gt; 118

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (19)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 118

Ser Leu Asp Ala Phe Arg Leu Ile Arg Ala Met Gly Ala Thr Gly Leu  
 1 5 10 15

Ser Phe Xaa

&lt;210&gt; 119

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

77

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (13)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 119

Leu Val Leu Trp Ile Val Met Leu Thr Tyr Ala Thr Xaa  
1 5 10

&lt;210&gt; 120

&lt;211&gt; 80

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (80)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 120

Met Glu Pro Gln Leu Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp  
1 5 10 15

Leu Ala Leu Leu Leu Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu  
20 25 30

Pro Ala Ser Ser Leu Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr  
35 40 45

Asn Phe Gly Arg Thr Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile  
50 55 60

Gly Thr Ser Ile Cys Lys Lys Phe Phe Lys Glu Arg Asn Lys Ile Xaa  
65 70 75 80

&lt;210&gt; 121

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (96)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (107)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (111)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (115)  
 <223> Xaa equals any of the naturally occurring L-amino acids  
  
 <220>  
 <221> SITE  
 <222> (122)  
 <223> Xaa equals any of the naturally occurring L-amino acids  
  
 <220>  
 <221> SITE  
 <222> (132)  
 <223> Xaa equals any of the naturally occurring L-amino acids  
  
 <400> 121  
 Met Glu Pro Gln Leu Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp  
   1                  5                  10                  15  
 Leu Ala Leu Leu Leu Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu  
                   20                  25                  30  
 Pro Ala Ser Ser Leu Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr  
           35                  40                  45  
 Asn Phe Gly Arg Thr Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile  
   50                  55                  60  
 Gly Thr Ser Ile Cys Lys Lys Phe Phe Lys Glu Glu Ile Arg Ser Asp  
   65                  70                  75                  80  
 Asn Trp Leu Ala Ser His Leu Gly Thr Ala Ser Arg Phe Pro Leu Xaa  
                   85                  90                  95  
 Ser Tyr Pro Cys Lys Leu Leu Gln Met Ile Xaa Lys Ile Trp Xaa Pro  
           100                  105                  110  
 Cys Gly Xaa Leu Leu Thr Gly Gln Gln Xaa Ser Asn Glu Ile Ser Lys  
   115                  120                  125  
 Gln Glu Ile Xaa Cys Leu Leu His Pro Pro Pro Lys Asn Leu His Ile  
   130                  135                  140  
 Asp Val  
 145  
  
 <210> 122  
 <211> 81  
 <212> PRT  
 <213> Homo sapiens  
  
 <220>  
 <221> SITE  
 <222> (81)  
 <223> Xaa equals stop translation  
  
 <400> 122

79

Met Ala Leu Gly Ser Met Tyr Leu Val Leu Thr Leu Ile Val Ala Lys  
 1                      5                      10                      15

Val Leu Arg Gly Ala Glu Pro Cys Cys Gly Pro Leu Lys Asn Arg Val  
                     20                      25                      30

Leu Arg Pro Cys Pro Leu Pro Val His Cys Pro Leu Pro Ile Pro Ser  
                     35                      40                      45

Pro Ala Glu Gly Ile Pro Trp Val Ala Tyr Leu Pro Ile Arg Trp Phe  
                     50                      55                      60

Ile Ser Cys Cys Pro Gly His Cys Ile Gln Ile Pro Met Cys Thr Ser  
                     65                      70                      75                      80

Xaa

<210> 123  
 <211> 337  
 <212> PRT  
 <213> Homo sapiens

<400> 123

Glu Pro His Arg Gly Pro His Leu Pro Pro Asp Leu Gly His His His  
 1                      5                      10                      15

Gly Gln Arg Pro Gly Leu Gln Asn Ile Asn Val Phe Leu Arg Asn Thr  
                     20                      25                      30

Val Lys Val Thr Gly Val Val Val Phe Met Phe Ser Leu Ser Trp Gln  
                     35                      40                      45

Leu Ser Leu Val Thr Phe Met Gly Phe Pro Ile Ile Met Met Val Ser  
                     50                      55                      60

Asn Ile Tyr Gly Lys Tyr Tyr Lys Arg Leu Ser Lys Glu Val Gln Asn  
                     65                      70                      75                      80

Ala Leu Ala Arg Ala Ser Asn Thr Ala Glu Glu Thr Ile Ser Ala Met  
                     85                      90                      95

Lys Thr Val Arg Ser Phe Ala Asn Glu Glu Glu Ala Glu Val Tyr  
                     100                      105                      110

Leu Arg Lys Leu Gln Gln Val Tyr Lys Leu Asn Arg Lys Glu Ala Ala  
                     115                      120                      125

Ala Tyr Met Tyr Tyr Val Trp Gly Ser Gly Leu Thr Leu Leu Val Val  
                     130                      135                      140

Gln Val Ser Ile Leu Tyr Tyr Gly Gly His Leu Val Ile Ser Gly Gln  
                     145                      150                      155                      160

Met Thr Ser Gly Asn Leu Ile Ala Phe Ile Ile Tyr Glu Phe Val Leu  
                     165                      170                      175

Gly Asp Cys Met Glu Asn Val Ser Phe Ser Leu Ser Pro Gly Lys Val

180 80 190  
 Thr Ala Leu Val Gly Pro Ser Gly Ser Gly Lys Ser Ser Cys Val Asn  
 195 200 205  
 Ile Leu Glu Asn Phe Tyr Pro Leu Glu Gly Gly Arg Val Leu Leu Asp  
 210 215 220  
 Gly Lys Pro Ile Ser Ala Tyr Asp His Lys Tyr Leu His Arg Val Ile  
 225 230 235 240  
 Ser Leu Val Ser Gln Glu Pro Val Leu Phe Ala Arg Ser Ile Thr Asp  
 245 250 255  
 Asn Ile Ser Tyr Gly Leu Pro Thr Val Pro Phe Glu Met Val Val Glu  
 260 265 270  
 Ala Ala Gln Lys Ala Asn Ala His Gly Phe Ile Met Glu Leu Gln Asp  
 275 280 285  
 Gly Tyr Ser Thr Glu Thr Gly Glu Lys Gly Ala Gln Leu Ser Gly Gly  
 290 295 300  
 Gln Lys Gln Arg Val Ala Trp Pro Gly Leu Trp Cys Gly Thr Pro Gln  
 305 310 315 320  
 Ser Ser Ser Trp Met Lys Pro Pro Ala Leu Trp Met Pro Arg Ala Ser  
 325 330 335  
 Ile

<210> 124  
 <211> 315  
 <212> PRT  
 <213> Homo sapiens

<400> 124  
 Met Ser Ser Ala Thr Trp Thr Ala Ala Ser Trp Arg Thr Ser Ala Thr  
 1 5 10 15  
 Ser Thr Ser Leu Thr Arg Cys Trp Ile Ser Gly Gln Pro Ala Cys Thr  
 20 25 30  
 Ala Ala Ala Cys Cys Trp Gly Ala Thr Ile Gly Val Ala Lys Asn Ser  
 35 40 45  
 Ala Leu Gly Pro Arg Arg Leu Arg Ala Ser Trp Leu Val Ile Thr Leu  
 50 55 60  
 Val Cys Leu Phe Val Gly Ile Tyr Ala Met Val Lys Leu Leu Leu Phe  
 65 70 75 80  
 Ser Glu Val Arg Arg Pro Ile Arg Asp Pro Trp Phe Trp Ala Leu Phe  
 85 90 95  
 Val Trp Thr Tyr Ile Ser Leu Gly Ala Ser Phe Leu Leu Trp Trp Leu  
 100 105 110

81

Leu Ser Thr Val Arg Pro Gly Thr Gln Ala Leu Glu Pro Gly Ala Ala  
 115 120 125  
 Thr Glu Ala Glu Gly Phe Pro Gly Ser Gly Arg Pro Pro Pro Glu Gln  
 130 135 140  
 Ala Ser Gly Ala Thr Leu Gln Lys Leu Leu Ser Tyr Thr Lys Pro Asp  
 145 150 155 160  
 Val Ala Phe Leu Val Ala Ala Ser Phe Phe Leu Ile Val Ala Ala Leu  
 165 170 175  
 Gly Glu Thr Phe Leu Pro Tyr Tyr Thr Gly Arg Ala Ile Asp Gly Ile  
 180 185 190  
 Val Ile Gln Lys Ser Met Asp Gln Phe Ser Thr Ala Val Val Ile Val  
 195 200 205  
 Cys Leu Leu Ala Ile Gly Ser Ser Phe Ala Ala Gly Ile Arg Gly Gly  
 210 215 220  
 Ile Phe Thr Leu Ile Phe Ala Arg Leu Asn Ile Arg Leu Arg Asn Cys  
 225 230 235 240  
 Leu Phe Arg Ser Leu Val Ser Gln Glu Thr Ser Phe Phe Asp Glu Asn  
 245 250 255  
 Arg Thr Gly Asp Leu Ile Ser Arg Leu Thr Ser Asp Thr Thr Met Val  
 260 265 270  
 Ser Asp Leu Val Ser Arg Thr Ser Met Ser Ser Cys Gly Thr Gln Ser  
 275 280 285  
 Arg Ser Arg Ala Trp Trp Ser Ser Cys Ser Ala Ser His Gly Ser Ser  
 290 295 300  
 Pro Trp Ser Pro Ser Trp Ala Ser Pro Ser Ser  
 305 310 315

&lt;210&gt; 125

&lt;211&gt; 167

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 125

His Leu Leu Arg Pro Ala His Cys Ala Phe Arg Asp Gly Gly Gly Gly  
 1 5 10 15  
 Arg Thr Glu Gly Gln Cys Pro Arg Leu His His Gly Thr Pro Gly Arg  
 20 25 30  
 Leu Gln His Arg Asp Arg Gly Glu Gly Arg Pro Ala Val Arg Trp Pro  
 35 40 45  
 Glu Ala Ala Gly Gly Met Ala Arg Ala Leu Val Arg Asn Pro Pro Val  
 50 55 60



```
<210> 126
<211> 227
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (71)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```

<400> 126
Arg Leu Thr Lys Thr Ile Ser Phe Ser Leu Gln Asn Gln Thr Ala Phe
  1              5              10              15

Ile Asn Ser Leu Ala Lys Thr Pro Tyr Gln Ala Leu Thr Gly Ala Ala
      20              25              30

Leu Ala Gly Ser Tyr Pro Ile Trp Glu Asn Glu Asn Thr Leu Ser Trp
      35              40              45

Tyr Leu Pro Ser Pro Thr Thr Leu Leu Ser Pro Pro Val Leu Phe Cys
      50              55              60

Val Ile Gln Leu Ile Phe Xaa Leu Pro Ala Asn Trp Ser Gly Thr Cys
      65              70              75              80

Thr Leu Val Phe Gln Ala Pro Thr Ile Asn Ile Leu Pro Pro Asn Gln
      85              90              95

Thr Ile Leu Ile Ser Val Glu Ala Ser Ile Ser Ser Ser Pro Ile Arg
      100              105              110

Asn Lys Trp Ala Leu His Leu Ile Thr Leu Leu Thr Gly Leu Gly Ile
      115              120              125

Thr Ala Ala Leu Gly Thr Gly Ile Ala Gly Ile Thr Thr Ser Ile Thr
      130              135              140

```

83

Ser Tyr Gln Thr Leu Phe Thr Thr Leu Ser Asn Thr Val Glu Asp Met  
145 150 155 160

His Thr Ser Ile Thr Ser Leu Gln Arg Gln Leu Asp Phe Leu Val Gly  
165 170 175

Val Ile Leu Gln Asn Trp Arg Val Leu Asp Leu Leu Thr Thr Glu Lys  
180 185 190

Gly Gly Thr Cys Ile Tyr Leu Gln Glu Glu Cys Cys Phe Cys Val Asn  
195 200 205

Glu Ser Gly Ile Val His Ile Ala Val Arg Arg Leu His Asp Arg Ala  
210 215 220

Ala Glu Leu  
225

<210> 127  
<211> 29  
<212> PRT  
<213> Homo sapiens

<400> 127  
Tyr Pro Ile Trp Glu Asn Glu Asn Thr Leu Ser Trp Tyr Leu Pro Ser  
1 5 10 15

Pro Thr Thr Leu Leu Ser Pro Pro Val Leu Phe Cys Val  
20 25

<210> 128  
<211> 27  
<212> PRT  
<213> Homo sapiens

<400> 128  
Arg Val Leu Asp Leu Leu Thr Thr Glu Lys Gly Gly Thr Cys Ile Tyr  
1 5 10 15

Leu Gln Glu Glu Cys Cys Phe Cys Val Asn Glu  
20 25

<210> 129  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 129  
Phe Ser Leu Gly Arg Arg His Cys Leu Gly  
1 5 10

<210> 130  
<211> 123  
<212> PRT

84

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (64)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (83)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;400&gt; 130

Glu	His	Pro	Thr	Ala	Asp	Arg	Ala	Gly	Cys	Ser	Ala	Ser	Gly	Ala	Cys
1				5				10					15		

Tyr	Ser	Leu	His	His	Ala	Thr	Met	Lys	Arg	Gln	Ala	Ala	Glu	Glu	Ala
		20						25					30		

Cys	Ile	Leu	Arg	Gly	Gly	Ala	Leu	Ser	Thr	Val	Arg	Ala	Gly	Ala	Glu
		35					40					45			

Leu	Arg	Ala	Val	Leu	Ala	Leu	Leu	Arg	Ala	Gly	Pro	Gly	Pro	Gly	Xaa
	50					55					60				

Gly	Ser	Lys	Asp	Leu	Leu	Phe	Trp	Val	Ala	Leu	Glu	Arg	Arg	Arg	Ser
65				70						75					80

His	Cys	Xaa	Leu	Glu	Asn	Glu	Pro	Leu	Arg	Gly	Phe	Ser	Trp	Leu	Ser
			85						90					95	

Ser	Asp	Pro	Gly	Gly	Leu	Glu	Ser	Asp	Thr	Leu	Gln	Trp	Val	Glu	Glu
		100						105					110		

Pro	Gln	Arg	Ser	Cys	Thr	Ala	Arg	Arg	Trp	Val
	115						120			

&lt;210&gt; 131

&lt;211&gt; 344

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (19)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;400&gt; 131

Ser	Arg	Pro	Pro	Val	Gly	Ser	Ser	Pro	Gln	Leu	Glu	Gly	Asp	Ala	Met
1				5					10				15		

Pro	Pro	Xaa	Arg	Gln	Arg	Tyr	Leu	Cys	Lys	Tyr	Gln	Phe	Glu	Val	Leu
		20						25					30		

Cys	Pro	Ala	Pro	Arg	Pro	Gly	Ala	Ala	Ser	Asn	Leu	Ser	Tyr	Arg	Ala
		35					40					45			

Pro	Phe	Gln	Leu	His	Ser	Ala	Ala	Leu	Asp	Phe	Ser	Pro	Pro	Gly	Thr
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

<210> 132

<211> 6

<212> PRT  
<213> Homo sapiens

<400> 132  
Val Glu Gly Gly His Ser  
1 5

<210> 133  
<211> 7  
<212> PRT  
<213> Homo sapiens

<400> 133  
Arg Tyr Leu Thr Leu Thr His  
1 5

<210> 134  
<211> 6  
<212> PRT  
<213> Homo sapiens

<400> 134  
Cys Asn Thr Pro Trp Ala  
1 5

<210> 135  
<211> 8  
<212> PRT  
<213> Homo sapiens

<400> 135  
Ala Pro Val Ile Phe Ser His Ser  
1 5

<210> 136  
<211> 6  
<212> PRT  
<213> Homo sapiens

<400> 136  
Arg Asn Val Pro Asp Asp  
1 5

<210> 137  
<211> 6  
<212> PRT  
<213> Homo sapiens

<400> 137  
Gly Leu Glu Asp Val Ser  
1 5

<210> 138

87

&lt;211&gt; 23

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 138

Val	Glu	Gly	Gly	His	Ser	Leu	Asp	Asn	Ser	Leu	Ser	Ile	Leu	Arg	Thr
1				5					10					15	

Phe	Tyr	Met	Leu	Gly	Val	Arg
			20			

&lt;210&gt; 139

&lt;211&gt; 190

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 139

Thr	Trp	Leu	Arg	Leu	Gly	Ser	Ser	Gln	Ile	Trp	Leu	Gly	Thr	Ala	Pro
1				5					10					15	

Arg	Gly	Pro	Arg	Ile	His	Pro	Glu	Gln	Ala	Gly	Leu	Ala	Gly	Ala	Pro
			20					25					30		

Val	Lys	Ser	Thr	Ser	Ser	Glu	Glu	Ser	Gln	Pro	Gly	Gly	Gln	Cys	Gln
	35						40					45			

Ser	Ser	Gly	Gly	Ala	Gln	Thr	Leu	Pro	Ser	Leu	Arg	Ala	Ala	Pro	Val
	50					55					60				

Ala	Ala	Leu	Gly	Ser	Leu	Ser	Ser	Tyr	Pro	Asp	Ser	Cys	Pro	Arg	Ala
65					70					75					80

Thr	Thr	Pro	Glu	Leu	Cys	Pro	Gly	Ala	Pro	Thr	Leu	His	Leu	Ala	Asp
				85					90					95	

Ser	Ile	Ser	Gly	Pro	Val	Ser	Pro	Pro	Gly	Ser	Ser	Leu	Gly	Pro	Asp
			100					105					110		

Ala	Trp	Thr	Leu	Cys	Ala	Lys	His	His	Gln	Ala	Lys	Gly	Met	Thr	Leu
		115					120					125			

Gly	Thr	Pro	Lys	Val	Leu	Arg	Leu	Gln	Pro	Val	Ser	Pro	Cys	Trp	Gly
	130					135					140				

Pro	Lys	Ser	Trp	Arg	Val	Pro	Gly	Pro	Phe	Gln	Pro	Gly	Arg	Arg	Arg
145					150					155					160

Gly	Glu	Ser	Arg	Gln	Gln	Gly	Arg	Gly	Lys	Arg	Arg	Ser	Ala	Arg	Ser
				165					170					175	

Ala	Gln	Ser	Pro	Thr	Gly	Pro	Glu	Ser	Ala	Ala	Trp	Pro	Cys
			180					185					190

&lt;210&gt; 140

&lt;211&gt; 129

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

88

&lt;400&gt; 140

Thr Val Ala Thr Ala Cys Val Trp Ala Ala Cys Thr Gly Cys Trp Ala  
 1 5 10 15

Arg Pro Pro Val Pro Thr Trp Ala Gly Cys Ala Ala Arg Cys Ala Ala  
 20 25 30

Glu Asp Ala Arg Ala Gly Val Gly Asp Leu Pro Ala Thr Gly Gly Ala  
 35 40 45

Ala Thr Gly Arg Arg Ala Leu Thr Pro Ala Pro Pro Arg Gly Pro Cys  
 50 55 60

Ile Leu Ser Pro Gln Pro Trp Ala Leu Gly Leu Pro Gly Ala Pro Leu  
 65 70 75 80

Pro Ala Ala Leu Pro Gly Arg Ala Arg Gly Arg Pro Gly Leu Pro Ala  
 85 90 95

Leu Pro Ala Leu Ser Thr Leu Pro Gly Cys Pro Ala Leu Asp Pro Ala  
 100 105 110

Gly Ala Gly Thr Leu Cys Pro Pro Pro Gly Ala Ala Glu Pro Ala Gly  
 115 120 125

Pro

&lt;210&gt; 141

&lt;211&gt; 90

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 141

Arg Ser Gly Gln Pro Gly Glu Gly Ser Met Leu Arg Lys Phe Ser Leu  
 1 5 10 15

Gln Arg Leu Leu Ser Pro Leu Asp Gln Ala Gln Thr Arg Trp Gly Leu  
 20 25 30

Ala Leu Ala Cys Val Ala Gly Asp Lys Gly Pro Pro Arg Pro Trp Asn  
 35 40 45

Ile Ser Ser Ala Pro Ala His Pro His Val Thr Thr Pro Gly Met Glu  
 50 55 60

Thr Ser Gly Gly Pro Ala Arg Asp Gly Gly Leu Ile Leu Glu Arg Glu  
 65 70 75 80

Ala Ala Phe Asn Lys Pro Ala Pro Gly Glu  
 85 90

&lt;210&gt; 142

&lt;211&gt; 307

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

<220>  
 <221> SITE  
 <222> (135)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (197)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (203)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (219)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (255)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 142  
 Arg Cys Gln Arg Asn Lys Asp Ile Met Met Ser Ser Lys Pro Thr Ser  
     1                    5                    10                    15  
 His Ala Glu Val Asn Glu Thr Ile Pro Asn Pro Tyr Pro Pro Ser Ser  
                     20                    25                    30  
 Phe Met Ala Pro Gly Phe Gln Gln Pro Leu Gly Ser Ile Asn Leu Glu  
             35                    40                    45  
 Asn Gln Ala Gln Gly Ala Gln Arg Ala Gln Pro Tyr Gly Ile Thr Ser  
     50                    55                    60  
 Pro Gly Ile Phe Ala Ser Ser Gln Pro Gly Gln Gly Asn Ile Gln Met  
     65                    70                    75                    80  
 Ile Asn Pro Ser Val Gly Thr Ala Val Met Asn Phe Lys Glu Glu Ala  
                     85                    90                    95  
 Lys Ala Leu Gly Val Ile Gln Ile Met Val Gly Leu Met His Ile Gly  
             100                    105                    110  
 Phe Gly Ile Val Leu Cys Leu Ile Ser Phe Ser Phe Arg Glu Val Leu  
     115                    120                    125  
 Gly Phe Ala Ser Thr Ala Xaa Ile Gly Gly Tyr Pro Phe Trp Gly Gly  
     130                    135                    140  
 Leu Ser Phe Ile Ile Ser Gly Ser Leu Ser Val Ser Ala Ser Lys Glu  
     145                    150                    155                    160  
 Leu Ser Arg Cys Leu Val Lys Gly Ser Leu Gly Met Asn Ile Gly Arg  
             165                    170                    175



90

Ser Ile Leu Ala Phe Ile Gly Val Ile Leu Leu Leu Val Asp Met Cys  
 180 185 190  
 Ile Asn Gly Val Xaa Gly Gln Asp Tyr Trp Xaa Val Leu Ser Gly Lys  
 195 200 205  
 Gly Ile Ser Ala Thr Leu Met Ile Phe Ser Xaa Leu Glu Phe Phe Val  
 210 215 220  
 Ala Cys Ala Thr Ala His Phe Ala Asn Gln Ala Asn Thr Thr Thr Asn  
 225 230 235 240  
 Met Ser Val Leu Val Ile Pro Asn Met Tyr Glu Ser Asn Pro Xaa Thr  
 245 250 255  
 Pro Ala Ser Ser Ser Ala Pro Pro Arg Cys Asn Asn Tyr Ser Ala Asn  
 260 265 270  
 Ala Pro Lys Arg Lys Arg Gly Ile Ser Leu Ile Ser Trp Arg Lys Thr  
 275 280 285  
 Thr Cys Lys Asn Phe Leu Arg Arg Cys Leu Leu Leu Ser Thr Met Ile  
 290 295 300  
 Ser Ser Leu  
 305

<210> 143  
 <211> 246  
 <212> PRT  
 <213> Homo sapiens

<400> 143  
 Met Gly Arg Leu Asp Gly Lys Val Ile Ile Leu Thr Ala Ala Ala Gln  
 1 5 10 15  
 Gly Ile Gly Gln Ala Ala Ala Leu Ala Phe Ala Arg Glu Gly Ala Lys  
 20 25 30  
 Val Ile Ala Thr Asp Ile Asn Glu Ser Lys Leu Gln Glu Leu Glu Lys  
 35 40 45  
 Tyr Pro Gly Ile Gln Thr Arg Val Leu Asp Val Thr Lys Lys Lys Gln  
 50 55 60  
 Ile Asp Gln Phe Ala Asn Glu Val Glu Arg Leu Asp Val Leu Phe Asn  
 65 70 75 80  
 Val Ala Gly Phe Val His His Gly Thr Val Leu Asp Cys Glu Glu Lys  
 85 90 95  
 Asp Trp Asp Phe Ser Met Asn Leu Asn Val Arg Asn Val Met Tyr Leu  
 100 105 110  
 Met Ile Lys Ala Phe Leu Pro Lys Met Leu Ala Gln Lys Ser Gly Asn  
 115 120 125

91

Ile Ile Asn Met Ser Ser Val Ala Ser Ser Val Lys Gly Val Val Asn  
 130 135 140

Arg Cys Val Tyr Ser Thr Thr Lys Ala Ala Val Ile Gly Leu Thr Lys  
 145 150 155 160

Ser Val Ala Ala Asp Phe Ile Gln Gln Gly Ile Arg Cys Asn Cys Val  
 165 170 175

Cys Pro Gly Thr Val Asp Thr Pro Ser Leu Gln Glu Arg Ile Gln Ala  
 180 185 190

Arg Gly Asn Pro Glu Glu Ala Arg Asn Asp Phe Leu Lys Arg Gln Lys  
 195 200 205

Thr Gly Arg Phe Ala Thr Ala Glu Glu Ile Ala Met Leu Cys Val Tyr  
 210 215 220

Leu Ala Ser Asp Glu Ser Ala Tyr Val Thr Gly Asn Pro Val Ile Ile  
 225 230 235 240

Asp Gly Gly Trp Ser Leu  
 245

&lt;210&gt; 144

&lt;211&gt; 234

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 144

Gly Thr Ile Gly Leu Tyr Trp Val Gly Ser Ile Ile Met Ser Val Val  
 1 5 10 15

Val Phe Val Pro Gly Asn Ile Val Gly Lys Tyr Gly Thr Arg Ile Cys  
 20 25 30

Pro Ala Phe Phe Leu Ser Ile Pro Tyr Thr Cys Leu Pro Val Trp Ala  
 35 40 45

Gly Phe Arg Ile Tyr Asn Gln Pro Ser Glu Asn Tyr Asn Tyr Pro Ser  
 50 55 60

Lys Val Ile Gln Glu Ala Gln Ala Lys Asp Leu Leu Arg Arg Pro Phe  
 65 70 75 80

Asp Leu Met Leu Val Val Cys Leu Leu Leu Ala Thr Gly Phe Cys Leu  
 85 90 95

Phe Arg Gly Leu Ile Ala Leu Asp Cys Pro Ser Glu Leu Cys Arg Leu  
 100 105 110

Tyr Thr Gln Phe Gln Glu Pro Tyr Leu Lys Asp Pro Ala Ala Tyr Pro  
 115 120 125

Lys Ile Gln Met Leu Ala Tyr Met Phe Tyr Ser Val Pro Tyr Phe Val  
 130 135 140

Thr Ala Leu Tyr Gly Leu Val Val Pro Gly Cys Ser Trp Met Pro Asp

```

<210> 145
<211> 238
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (184)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 145
Met Ser Asn His Asp Pro Arg Gly Cys Thr Arg Arg Arg Ala Gln Lys
  1              5              10              15

Pro Leu Ala Ile Gln Pro Arg Leu Phe His Ala Ser Ala Pro Asp Glu
      20              25              30

Gly Thr Gln Gly Thr Leu Lys Gly Thr Gln Lys Gly Gly Cys Ile Leu
      35              40              45

Val Gln Cys Gln Ser Glu Gly Gly Ala Ala Gly Ala Trp Thr Gly Pro
      50              55              60

Pro Ser Pro Ala Arg Asp Arg Arg Val Arg Pro Pro Gly Thr Lys Ala
      65              70              75              80

Gln Arg Leu Glu Arg Arg Arg His Val Pro Arg Leu His Gly Leu Gly
      85              90              95

Val Gly Gly Cys Glu Val Arg Thr Gly Ile Val Ala Arg Ile Ser Gly
      100             105             110

Ser Thr Pro Trp Ala Gly Gly Lys Pro Leu Gly Leu His Gly Ala Met
      115             120             125

Gly Glu Ala Gly Ala Gly Asp Thr Gly Cys Cys Ala Lys Gly Pro Ser
      130             135             140

Pro Ala Ala Pro Leu Pro Ala Glu Gly Arg Gly Gln Gly Ala Gly Pro
145             150             155             160

```

Ala Ala Leu Ala Cys Ser Ala Gly Asp Thr Trp Ala Ser Glu Val Gly  
50 55 60

94

Pro Val Leu Ser Lys Ser Ser Pro Arg Leu Ile Thr Thr Trp Glu Lys  
 65                      70                      75                      80

Val Pro Val Gly Thr Asn Gly Gly Val Thr Val Val Gly Leu Val Ser  
                             85                      90                      95

Ser Leu Leu Gly Gly Thr Phe Val Gly Ile Ala Tyr Phe Leu Thr Gln  
                             100                      105                      110

Leu Ile Phe Val Asn Asp Leu Asp Ile Ser Ala Pro Gln Trp Pro Ile  
                             115                      120                      125

Ile Ala Phe Gly Gly Leu Ala Gly Leu Leu Gly Ser Ile Val Asp Ser  
                             130                      135                      140

Tyr Leu Gly Ala Thr Met Gln Tyr Thr Gly Leu Asp Glu Ser Thr Gly  
 145                      150                      155                      160

Met Val Val Asn Ser Pro Thr Asn Xaa Ala Arg His Ile Ala Gly Lys  
                             165                      170                      175

Pro Ile Leu Asp Asn Asn Ala Val Asn Leu Phe Ser Ser Val Leu Ile  
                             180                      185                      190

Ala Leu Leu Leu Pro Thr Ala Ala Trp Gly Phe Trp Pro Arg Gly  
                             195                      200                      205

&lt;210&gt; 147

&lt;211&gt; 116

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 147

Met Ser Gln Arg Ala Gly Arg Arg Pro Gly Gly Trp Asn Pro Ser Leu  
 1                      5                      10                      15

Ser Val Val Glu Val Cys Arg Gly Cys Arg Gly Thr Gly Pro Leu Pro  
                             20                      25                      30

Trp Gly Ala Ser Leu Phe Pro Cys Ser Ala Ser Pro Leu Phe Pro Leu  
                             35                      40                      45

Pro Leu Asn Arg Arg Gly Asp Val His Gly Thr Leu Gly Gly Arg Met  
                             50                      55                      60

Leu Asn Arg Val Glu Cys Arg Asp Gly Val Ala Ala Ala Trp Leu Cys  
                             65                      70                      75                      80

Leu His Asp Ala Ala Ala Ile Arg Gly Ala Val Gly Arg Cys Pro Met  
                             85                      90                      95

Trp Thr Gln Pro Thr His Trp Val Leu Leu Leu Cys Trp Ala Leu His  
                             100                      105                      110

Phe Tyr Cys Arg  
                             115

95

&lt;210&gt; 148

&lt;211&gt; 81

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 148

Met Thr Ala His Ser Phe Ala Leu Pro Val Ile Ile Phe Thr Thr Phe  
 1 5 10 15

Trp Gly Leu Val Gly Ile Ala Gly Pro Trp Phe Val Pro Lys Gly Pro  
 20 25 30

Asn Arg Gly Val Ile Ile Thr Met Leu Val Ala Thr Ala Val Cys Cys  
 35 40 45

Tyr Leu Phe Trp Leu Ile Ala Ile Leu Ala Gln Leu Asn Pro Leu Phe  
 50 55 60

Gly Pro Gln Leu Lys Asn Glu Thr Ile Trp Tyr Val Arg Phe Leu Trp  
 65 70 75 80

Glu

&lt;210&gt; 149

&lt;211&gt; 110

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 149

Ala Gln Arg Ala Ala Arg Leu Gly Thr Arg Ala Pro Ala Ala Pro Ala  
 1 5 10 15

Ala Arg Pro Cys Ile Leu Pro Gly His Pro Ala Pro Gly His Asp Gly  
 20 25 30

Ala Leu Ile Arg Pro Pro Gly His His Leu His His Val Leu Gly Pro  
 35 40 45

Arg Arg His Arg Gly Pro Trp Phe Val Pro Lys Gly Pro Asn Arg Gly  
 50 55 60

Val Ile Ile Thr Met Leu Val Ala Thr Ala Val Cys Cys Tyr Leu Phe  
 65 70 75 80

Trp Leu Ile Ala Ile Leu Ala Gln Leu Asn Pro Leu Phe Gly Pro Gln  
 85 90 95

Leu Lys Asn Glu Thr Ile Trp Tyr Val Arg Phe Leu Trp Glu  
 100 105 110

&lt;210&gt; 150

&lt;211&gt; 135

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 150

96

Met	Thr	Leu	Glu	Glu	His	Arg	Asp	Arg	Pro	Arg	Leu	Gly	Met	Cys	Met
1				5					10					15	
Cys	Val	Cys	Ala	Cys	Val	Tyr	Ala	Cys	Met	Leu	Met	His	Val	Cys	Val
			20					25					30		
His	Ala	Cys	Leu	Cys	Val	Cys	Val	Cys	Val	Cys	Val	Glu	Pro	Trp	Ser
		35					40					45			
Ser	Arg	Gln	Ser	Lys	Asp	Thr	Gly	Gly	Trp	His	Met	Glu	Glu	Gln	Val
	50					55					60				
Thr	Pro	Pro	Ser	Leu	Ala	Gln	Leu	Lys	Ser	Gly	Gln	Val	Arg	Gly	Glu
65					70					75					80
Met	Gly	Glu	Gly	Arg	Gly	Glu	Lys	Gly	Glu	Glu	Ala	Leu	Thr	Gly	Gly
				85					90					95	
Ala	Glu	Ala	Leu	Ser	Leu	Leu	Gly	Arg	Arg	Ser	Pro	Ser	Thr	Pro	Leu
			100					105					110		
Phe	Leu	Asp	Arg	Glu	Asp	Lys	Gln	Ala	Lys	Asp	Ala	Arg	Asn	Leu	Ser
		115					120					125			
Ser	Thr	Val	Ala	Pro	Asp	Phe									
	130					135									

<210> 151  
 <211> 82  
 <212> PRT  
 <213> Homo sapiens

<400> 151

His	Glu	Lys	Ile	Leu	Thr	Pro	Ile	Trp	Pro	Ser	Ser	Thr	Asp	Leu	Glu
1				5					10					15	
Lys	Pro	His	Glu	Met	Leu	Phe	Leu	Asn	Val	Ile	Leu	Phe	Ser	Leu	Thr
			20					25					30		
Val	Phe	Thr	Leu	Ile	Ser	Thr	Ala	His	Thr	Leu	Asp	Arg	Ala	Val	Arg
		35					40					45			
Ser	Asp	Trp	Leu	Leu	Leu	Val	Leu	Ile	Tyr	Ala	Cys	Leu	Glu	Glu	Leu
	50					55					60				
Ile	Pro	Glu	Leu	Ile	Phe	Asn	Leu	Tyr	Cys	Gln	Gly	Asn	Ala	Thr	Leu
65					70					75				80	

Phe Phe

<210> 152  
 <211> 71  
 <212> PRT  
 <213> Homo sapiens

<400> 152

97

Pro Ala Asn Lys Ala Gly Ala Ala Ile Glu Ala Gly Ile Gly Ile Ser  
 1 5 10 15

Leu Met Val Leu Ser Pro Trp Ala Cys Leu Phe Val Val Phe Phe Pro  
 20 25 30

Tyr Ile Gln Ser Ser Leu Arg Ser Asp Lys His Leu Gln Leu Ser Asn  
 35 40 45

Ile Leu Pro Thr Pro Ser His His Ile His Leu Pro Ala Ser Ile Cys  
 50 55 60

Ile Gln Leu Arg Ala Gly Asn  
 65 70

&lt;210&gt; 153

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 153

Ala Gly Ser Pro Ala Gly Thr Gly Pro Glu Phe Pro Gly Arg Pro Thr  
 1 5 10 15

Arg Pro Ile Ser Thr His Val Phe Glu Tyr Glu Cys Ile Cys Lys Ile  
 20 25 30

Pro Arg Phe Met Cys Glu Tyr Val Leu Leu Leu Tyr Ile Val Leu Leu  
 35 40 45

Cys Asn Arg Ser Tyr Ala Val Phe Thr Gln Cys Val Leu Arg Ser Ser  
 50 55 60

Pro Ile Asp Ser Ser Arg Asn Ala Val Leu Leu  
 65 70 75

&lt;210&gt; 154

&lt;211&gt; 483

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (194)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (205)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;400&gt; 154

Met Pro Ser Gly Met Ser Ala Ala Val Pro Ile Ser Gly Leu Leu Asp  
 1 5 10 15

Leu Ser His Asn Ser Ile Ser Gln Glu Ser Ala Leu Tyr Leu Leu Glu  
 20 25 30



Thr Leu Pro Ser Cys Pro Arg Val Arg Glu Ala Ser Val Asn Leu Gly  
           35                          40                          45  
 Ser Glu Gln Ser Phe Arg Ile His Phe Ser Arg Glu Asp Gln Ala Gly  
           50                          55                          60  
 Lys Thr Leu Arg Leu Ser Glu Cys Ser Phe Arg Pro Glu His Val Ser  
           65                          70                          75                          80  
 Arg Leu Ala Thr Gly Leu Ser Lys Ser Leu Gln Leu Thr Glu Leu Thr  
                           85                          90                          95  
 Leu Thr Gln Cys Cys Leu Gly Gln Lys Gln Leu Ala Ile Leu Leu Ser  
                           100                          105                          110  
 Leu Val Gly Arg Pro Ala Gly Leu Phe Ser Leu Arg Val Gln Glu Pro  
           115                          120                          125  
 Trp Ala Asp Arg Ala Arg Val Leu Ser Leu Leu Glu Val Cys Ala Gln  
           130                          135                          140  
 Ala Ser Gly Ser Val Thr Glu Ile Ser Ile Ser Glu Thr Gln Gln Gln  
           145                          150                          155                          160  
 Leu Cys Val Gln Leu Glu Phe Pro Arg Gln Glu Glu Asn Pro Glu Ala  
                           165                          170                          175  
 Val Ala Leu Arg Leu Ala His Cys Asp Leu Gly Ala His His Ser Leu  
           180                          185                          190  
 Leu Xaa Gly Gln Leu Met Glu Thr Cys Ala Arg Leu Xaa Gln Leu Ser  
           195                          200                          205  
 Leu Ser Gln Val Asn Leu Cys Glu Asp Asp Asp Ala Ser Ser Leu Leu  
           210                          215                          220  
 Leu Gln Ser Leu Leu Leu Ser Leu Ser Glu Leu Lys Thr Phe Arg Leu  
           225                          230                          235                          240  
 Thr Ser Ser Cys Val Ser Thr Glu Gly Leu Ala His Leu Ala Ser Gly  
                           245                          250                          255  
 Leu Gly His Cys His His Leu Glu Glu Leu Asp Leu Ser Asn Asn Gln  
           260                          265                          270  
 Phe Asp Glu Glu Gly Thr Lys Ala Leu Met Arg Ala Leu Glu Gly Lys  
           275                          280                          285  
 Trp Met Leu Lys Arg Leu Asp Leu Ser His Leu Leu Leu Asn Ser Ser  
           290                          295                          300  
 Thr Leu Ala Leu Leu Thr His Arg Leu Ser Gln Met Thr Cys Leu Gln  
           305                          310                          315                          320  
 Ser Leu Arg Leu Asn Arg Asn Ser Ile Gly Asp Val Gly Cys Cys His  
           325                          330                          335  
 Leu Ser Glu Ala Leu Arg Ala Ala Thr Ser Leu Glu Glu Leu Asp Leu

99

340 345 350  
 Ser His Asn Gln Ile Gly Asp Ala Gly Val Gln His Leu Ala Thr Ile  
 355 360 365  
 Leu Pro Gly Leu Pro Glu Leu Arg Lys Ile Asp Leu Ser Gly Asn Ser  
 370 375 380  
 Ile Ser Ser Ala Gly Gly Val Gln Leu Ala Glu Ser Leu Val Leu Cys  
 385 390 395 400  
 Arg Arg Leu Glu Glu Leu Met Leu Gly Cys Asn Ala Leu Gly Asp Pro  
 405 410 415  
 Thr Ala Leu Gly Leu Ala Gln Glu Leu Pro Gln His Leu Arg Val Leu  
 420 425 430  
 His Leu Pro Phe Ser His Leu Gly Pro Gly Gly Ala Leu Ser Leu Ala  
 435 440 445  
 Arg Pro Trp Met Asp Pro Pro Ile Trp Lys Arg Ser Ala Trp Arg Lys  
 450 455 460  
 Thr Thr Trp Leu Glu Gly Ser Cys Val Ser Val Trp Ser Ser Arg Cys  
 465 470 475 480  
 Ser Asp Arg

&lt;210&gt; 155

&lt;211&gt; 221

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 155

His Gln Leu Ser Arg Gly Ser Ala Val Gly Arg Val Ser Arg Ser Leu  
 1 5 10 15  
 Gln Ala Pro Gly Gly Val Asp Ala Trp Leu Gln Cys Pro Gly Gly Ser  
 20 25 30  
 His Ser Pro Gly Ala Gly Ser Gly Ala Ala Pro Ala Pro Glu Gly Pro  
 35 40 45  
 Thr Pro Thr Ile Gln Pro Ser Gly Pro Arg Trp Gly Pro Glu Pro Gly  
 50 55 60  
 Gln Ala Leu Asp Gly Ser Pro His Leu Glu Glu Ile Ser Leu Ala Glu  
 65 70 75 80  
 Asn Asn Leu Ala Gly Gly Val Leu Arg Phe Cys Met Glu Leu Pro Leu  
 85 90 95  
 Leu Arg Gln Ile Asp Leu Val Ser Cys Lys Ile Asp Asn Gln Thr Ala  
 100 105 110  
 Lys Leu Leu Thr Ser Ser Phe Thr Ser Cys Pro Ala Leu Glu Val Ile  
 115 120 125

100

Leu Leu Ser Trp Asn Leu Leu Gly Asp Glu Ala Ala Ala Glu Leu Ala  
 130 135 140  
 Gln Val Leu Pro Gln Met Gly Arg Leu Lys Arg Val Asp Leu Glu Lys  
 145 150 155 160  
 Asn Gln Ile Thr Ala Leu Gly Ala Trp Leu Leu Ala Glu Gly Leu Ala  
 165 170 175  
 Gln Gly Ser Ser Ile Gln Val Ile Arg Leu Trp Asn Asn Pro Ile Pro  
 180 185 190  
 Cys Asp Met Ala Gln His Leu Lys Ser Gln Glu Pro Arg Leu Asp Phe  
 195 200 205  
 Ala Phe Phe Asp Asn Gln Pro Gln Ala Pro Trp Gly Thr  
 210 215 220

<210> 156  
 <211> 89  
 <212> PRT  
 <213> Homo sapiens

<400> 156  
 Glu Lys Leu Phe Cys Phe Glu Met Leu Leu Ile Cys Lys Phe Ser Pro  
 1 5 10 15  
 Asn Ser Val Pro Pro Glu Thr Cys Ala Ile Leu Asn Gln Gly Leu Met  
 20 25 30  
 Asp Leu Gly Leu Cys Arg Met Cys Leu Gly Asn Asn Met Phe Ala Gly  
 35 40 45  
 Ser Met Leu Gly Lys Ser His Arg His Ser Pro Phe Ser Ile Asn Gln  
 50 55 60  
 Arg His Asn Ala Leu Arg Lys Ala Ala Gly Thr Pro Ala Gln Lys Ser  
 65 70 75 80  
 Leu Gly Ile Val Gln Val Ser Pro Asn  
 85

<210> 157  
 <211> 58  
 <212> PRT  
 <213> Homo sapiens

<400> 157  
 Gly Cys Ala Gly Cys Ala Leu Val Thr Ile Cys Leu Gln Ala Val Cys  
 1 5 10 15  
 Leu Val Lys Ala Ile Ala Ile Leu His Ser Arg Leu Thr Arg Asp Thr  
 20 25 30  
 Met His Cys Gly Arg Pro Gln Gly Pro Leu Pro Arg Lys Ala Trp Val  
 35 40 45

101

Leu Ser Arg Phe Pro Pro Thr Glu Thr Ala  
50 55

<210> 158  
<211> 48  
<212> PRT  
<213> Homo sapiens

<400> 158  
Pro Glu Thr Gln Cys Thr Ala Glu Gly Arg Arg Asp Pro Cys Pro Glu  
1 5 10 15  
Lys Pro Gly Tyr Cys Pro Gly Phe Pro Gln Leu Arg Gln Pro Glu Ile  
20 25 30  
Trp Pro Arg Gly Lys Gly Lys Thr Leu His Pro Pro Ala Arg His Met  
35 40 45

<210> 159  
<211> 8  
<212> PRT  
<213> Homo sapiens

<400> 159  
Ser Glu Ile Gly Glu Asn Arg Pro  
1 5

<210> 160  
<211> 8  
<212> PRT  
<213> Homo sapiens

<400> 160  
His Asp Thr Asp Ser Phe Ala His  
1 5

<210> 161  
<211> 7  
<212> PRT  
<213> Homo sapiens

<400> 161  
Ala Leu Arg Lys Ala Ala Gly  
1 5

<210> 162  
<211> 148  
<212> PRT  
<213> Homo sapiens

102

&lt;400&gt; 162

Met Arg Gly Pro Val Cys Gly Phe Ser Leu Val Glu Met Leu Leu Ala  
 1 5 10 15

Leu Ala Leu Gly Leu Met Leu Ile Leu Gly Val Thr Gln Ile Ala Leu  
 20 25 30

Ser Ser Arg Thr Thr Tyr Ala Ser Gln Ser Ala Ala Ser Leu Leu Gln  
 35 40 45

Asp Asp Ala Arg Phe Ala Leu Gly Lys Leu Ile Gln Glu Ile Arg Gln  
 50 55 60

Ala Gly Met Phe Gly Cys Leu Ser Ala Ala Ser Ile Ser Asn Ala Pro  
 65 70 75 80

Ala Gly Phe Asp Arg Pro Ile Gly Trp Ser Thr Thr Gly Ser Ser Arg  
 85 90 95

Ser Leu Thr Leu Val Thr Ala Asp Val Gly Glu Gly Gly Ser Lys Pro  
 100 105 110

Asp Trp Thr Val Leu Ser Asp Cys Thr Gly Ser Ala His Ala Tyr Val  
 115 120 125

Gly Ser Pro Pro Ala Ala Asn Ala Arg Ala Asn Pro Leu Pro Thr Cys  
 130 135 140

Ala Lys Leu Thr  
 145

&lt;210&gt; 163

&lt;211&gt; 137

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 163

Met Gly Tyr Tyr Leu Ser Arg Ser Arg Gln Ala Gly Met Val Leu Leu  
 1 5 10 15

Ile Ser Leu Val Phe Leu Leu Leu Leu Ala Leu Leu Gly Val Ser Ser  
 20 25 30

Met Gln Gly Ala Ile Ser Gln Glu Lys Ile Thr Gly Ser Leu Arg Gln  
 35 40 45

Arg Asn Gln Ser Phe Gln Gln Ala Glu Ser Gly Leu Arg Leu Gly Glu  
 50 55 60

Ser Leu Val Gln Ala Ser Gly Phe Ala Leu Arg Pro Cys His Ser Thr  
 65 70 75 80

Ala Ala Cys Ala Pro Pro Ala Glu Ser Val Ser Val Val Gly Pro Gly  
 85 90 95

Thr Asn Pro Val Ser Thr Val Thr Trp Ile Gly Met Lys Asp Gly Val  
 100 105 110

103

Tyr Gly Ile Gln Asn Leu Gly Pro Gly Thr Gly Leu Val Asn Ser Arg  
 115 120 125

Gln Arg Pro Arg Pro Arg Ser Ile Ala  
 130 135

&lt;210&gt; 164

&lt;211&gt; 209

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 164

Glu Asn Glu Ser Thr Lys Glu Pro Ser Leu Leu Gln Tyr Leu Cys Val  
 1 5 10 15

Gln Ser Pro Ala Gly Leu Asn Gly Phe Asn Val Leu Leu Ser Gly Ser  
 20 25 30

Gln Thr Pro Pro Thr Val Gly Pro Ser Ser Gly Gln Leu Pro Ser Phe  
 35 40 45

Ser Val Pro Cys Met Val Leu Pro Ser Pro Pro Leu Gly Pro Phe Pro  
 50 55 60

Val Leu Tyr Ser Pro Ala Met Pro Gly Pro Val Ser Ser Thr Leu Gly  
 65 70 75 80

Ala Leu Pro Asn Thr Gly Pro Val Asn Phe Ser Leu Pro Gly Leu Gly  
 85 90 95

Ser Ile Ala Gln Leu Leu Val Gly Pro Thr Ala Val Val Asn Pro Lys  
 100 105 110

Ser Ser Thr Leu Pro Ser Ala Asp Pro Gln Leu Gln Ser Gln Pro Ser  
 115 120 125

Leu Asn Leu Ser Pro Val Met Ser Arg Ser His Ser Val Val Gln Gln  
 130 135 140

Pro Glu Ser Pro Val Tyr Val Gly His Pro Val Ser Val Val Lys Leu  
 145 150 155 160

His Gln Ser Pro Val Pro Val Thr Pro Lys Ser Ile Gln Arg Thr His  
 165 170 175

Arg Glu Thr Phe Phe Lys Thr Pro Gly Ser Leu Gly Asp Pro Val Leu  
 180 185 190

Lys Arg Arg Glu Arg Asn Asn His Glu Thr Pro Ala Arg Pro Arg Gly  
 195 200 205

Asp

&lt;210&gt; 165

&lt;211&gt; 454

&lt;212&gt; PRT

104

&lt;213&gt; Homo sapiens

&lt;400&gt; 165

Arg His Glu Arg His Glu Tyr Arg Arg Ala Leu Asp His Glu Glu Glu  
 1 5 10 15  
 Ala Leu Ser Ser Gly Ser Val Gln Glu Ala Glu Ala Met Leu Asp Glu  
 20 25 30  
 Pro Gln Glu Gln Ala Glu Gly Ser Leu Thr Val Tyr Val Ile Ser Glu  
 35 40 45  
 His Ser Ser Leu Leu Pro Gln Asp Met Met Ser Tyr Ile Gly Pro Lys  
 50 55 60  
 Arg Thr Ala Val Val Arg Gly Ile Met His Arg Glu Ala Phe Asn Ile  
 65 70 75 80  
 Ile Gly Arg Arg Ile Val Gln Val Ala Gln Ala Met Ser Leu Thr Glu  
 85 90 95  
 Asp Val Leu Ala Ala Ala Leu Ala Asp His Leu Pro Glu Asp Lys Trp  
 100 105 110  
 Ser Ala Glu Lys Arg Arg Pro Leu Lys Ser Ser Leu Gly Tyr Glu Ile  
 115 120 125  
 Thr Phe Ser Leu Leu Asn Pro Asp Pro Lys Ser His Asp Val Tyr Trp  
 130 135 140  
 Asp Ile Glu Gly Ala Val Arg Arg Tyr Val Gln Pro Phe Leu Asn Ala  
 145 150 155 160  
 Leu Gly Ala Ala Gly Asn Phe Ser Val Asp Ser Gln Ile Leu Tyr Tyr  
 165 170 175  
 Ala Met Leu Gly Val Asn Pro Arg Phe Asp Ser Ala Ser Ser Ser Tyr  
 180 185 190  
 Tyr Leu Asp Met His Ser Leu Pro His Val Ile Asn Pro Val Glu Ser  
 195 200 205  
 Arg Leu Gly Ser Ser Ala Ala Ser Leu Tyr Pro Val Leu Asn Phe Leu  
 210 215 220  
 Leu Tyr Val Pro Glu Leu Ala His Ser Pro Leu Tyr Ile Gln Asp Lys  
 225 230 235 240  
 Asp Gly Ala Pro Val Ala Thr Asn Ala Phe His Ser Pro Arg Trp Gly  
 245 250 255  
 Gly Ile Met Val Tyr Asn Val Asp Ser Lys Thr Tyr Asn Ala Ser Val  
 260 265 270  
 Leu Pro Val Arg Val Glu Val Asp Met Val Arg Val Met Glu Val Phe  
 275 280 285  
 Leu Ala Gln Leu Arg Leu Leu Phe Gly Ile Ala Gln Pro Gln Leu Pro  
 290 295 300

105

Pro Lys Cys Leu Leu Ser Gly Pro Thr Ser Glu Gly Leu Met Thr Trp  
 305 310 315 320  
 Glu Leu Asp Arg Leu Leu Trp Ala Arg Ser Val Glu Asn Leu Ala Thr  
 325 330 335  
 Ala Thr Thr Thr Leu Thr Ser Leu Ala Gln Leu Leu Gly Lys Ile Ser  
 340 345 350  
 Asn Ile Val Ile Lys Asp Asp Val Ala Ser Glu Val Tyr Lys Ala Val  
 355 360 365  
 Ala Ala Val Gln Lys Ser Ala Glu Glu Leu Ala Ser Gly His Leu Ala  
 370 375 380  
 Ser Ala Phe Val Ala Ser Gln Glu Ala Val Thr Ser Ser Glu Leu Ala  
 385 390 395 400  
 Phe Phe Asp Pro Ser Leu Leu His Leu Leu Tyr Phe Pro Asp Asp Gln  
 405 410 415  
 Lys Phe Ala Ile Tyr Ile Pro Leu Phe Leu Pro Met Ala Val Pro Ile  
 420 425 430  
 Leu Leu Ser Leu Val Lys Ile Phe Leu Glu Thr Arg Lys Ser Trp Arg  
 435 440 445  
 Lys Pro Glu Lys Thr Asp  
 450

<210> 166  
 <211> 66  
 <212> PRT  
 <213> Homo sapiens

<400> 166  
 Lys Leu Leu Leu Thr Lys Val Glu Gln Lys Leu Glu Leu Ala Arg Leu  
 1 5 10 15  
 Gln Val Asp Thr Ser Gly Ser Lys Glu Phe Gly Thr Ser Gly Ile Pro  
 20 25 30  
 Ala Lys Cys Arg Phe Pro Lys Ile Phe Val Asn Thr Asp Asp Thr Tyr  
 35 40 45  
 Glu Glu Leu His Leu Ile Val Tyr Lys Val Thr Thr Val Phe Leu Pro  
 50 55 60  
 Ala Leu  
 65

<210> 167  
 <211> 79  
 <212> PRT  
 <213> Homo sapiens



106

&lt;400&gt; 167

Met Glu Pro Gln Leu Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp  
 1 5 10 15  
 Leu Ala Leu Leu Leu Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu  
 20 25 30  
 Pro Ala Ser Ser Leu Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr  
 35 40 45  
 Asn Phe Gly Arg Thr Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile  
 50 55 60  
 Gly Thr Ser Ile Cys Lys Lys Phe Phe Lys Glu Arg Asn Lys Ile  
 65 70 75

&lt;210&gt; 168

&lt;211&gt; 209

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 168

Gln Leu Pro Leu Trp Pro Ser Pro Ala Ser Val Gln Pro Arg Val Asp  
 1 5 10 15  
 Ser Gln Arg Ala Arg Gly Ser Pro Glu Pro Lys Met Glu Pro Gln Leu  
 20 25 30  
 Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp Leu Ala Leu Leu Leu  
 35 40 45  
 Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu Pro Ala Ser Ser Leu  
 50 55 60  
 Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr Asn Phe Gly Arg Thr  
 65 70 75 80  
 Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile Gly Thr Ser Ile Cys  
 85 90 95  
 Lys Lys Phe Phe Lys Glu Glu Ile Arg Ser Asp Asn Trp Leu Ala Ser  
 100 105 110  
 His Leu Gly Leu Pro Pro Asp Ser Leu Leu Ser Tyr Pro Ala Asn Tyr  
 115 120 125  
 Ser Asp Asp Ser Lys Ile Trp Arg Pro Val Glu Ile Phe Arg Leu Val  
 130 135 140  
 Ser Lys Tyr Gln Asn Glu Ile Ser Asp Arg Lys Ile Cys Ala Ser Ala  
 145 150 155 160  
 Ser Ala Pro Lys Thr Cys Ser Ile Glu Arg Val Leu Arg Lys Thr Glu  
 165 170 175  
 Arg Phe Gln Lys Trp Leu Gln Ala Lys Arg Leu Thr Pro Asp Leu Val  
 180 185 190

107

Gln Asp Cys His Gln Gly Gln Arg Glu Leu Lys Phe Leu Cys Met Leu  
 195 200 205

Arg

<210> 169  
 <211> 146  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (96)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (107)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (111)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (115)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (122)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (132)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 169  
 Met Glu Pro Gln Leu Gly Pro Glu Ala Ala Ala Leu Arg Pro Gly Trp  
 1 5 10 15

Leu Ala Leu Leu Leu Trp Val Ser Ala Leu Ser Cys Ser Phe Ser Leu  
 20 25 30

Pro Ala Ser Ser Leu Ser Ser Leu Val Pro Gln Val Arg Thr Ser Tyr  
 35 40 45

Asn Phe Gly Arg Thr Phe Leu Gly Leu Asp Lys Cys Asn Ala Cys Ile  
 50 55 60

Gly Thr Ser Ile Cys Lys Lys Phe Phe Lys Glu Glu Ile Arg Ser Asp  
 65 70 75 80

Asn Trp Leu Ala Ser His Leu Gly Thr Ala Ser Arg Phe Pro Leu Xaa  
 85 90 95

108

Ser Tyr Pro Cys Lys Leu Leu Gln Met Ile Xaa Lys Ile Trp Xaa Pro  
 100 105 110

Cys Gly Xaa Leu Leu Thr Gly Gln Gln Xaa Ser Asn Glu Ile Ser Lys  
 115 120 125

Gln Glu Ile Xaa Cys Leu Leu His Pro Pro Pro Lys Asn Leu His Ile  
 130 135 140

Asp Val  
 145

<210> 170

<211> 97

<212> PRT

<213> Homo sapiens

<400> 170

Gly Pro Arg Ala Arg Val Gln Gly Phe Ser Gly Ala Asp Ile Val Lys  
 1 5 10 15

Phe Met Ala Leu Gly Ser Met Tyr Leu Val Leu Thr Leu Ile Val Ala  
 20 25 30

Lys Val Leu Arg Gly Ala Glu Pro Cys Cys Gly Pro Leu Lys Asn Arg  
 35 40 45

Val Leu Arg Pro Cys Pro Leu Pro Val His Cys Pro Leu Pro Ile Pro  
 50 55 60

Ser Pro Ala Glu Gly Ile Pro Trp Val Ala Tyr Leu Pro Ile Arg Trp  
 65 70 75 80

Phe Ile Ser Cys Cys Pro Gly His Cys Ile Gln Ile Pro Met Cys Thr  
 85 90 95

Ser

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19330

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) : C07H 21/04; C07K 1/00; C12N 1/20; C12P 21/06

US CL : 530/350; 536/23.5; 435/252.3; 435/69.1

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.5; 435/252.3; 435/69.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Sequence databases: emb158, genbank111, n-issued, n-geneseq35, emb1-c3st58, agenseq35, a-issued, pir60

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genebank, accession U66676, Homo sapiens cDNA similar to ATP-binding cassette transporter. 12 September 1996, Vol. 5, No. 10, pages 1649-1655, see attached sequence alignment.	1-10, 21
X	Database Genbank, accession N46602, Homo sapiens cDNA clone IMAGE: 276822, April 1993, see attached sequence alignment.	1-10, 21
Y	US 5,352,596 A (CHEUNG et al) 04 October 1994, col 35, line 1, and see attached sequence alignment.	11,12, 14-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

01 DECEMBER 1999

Date of mailing of the international search report

7 fevrier 2000 07/02/2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

MICHAEL T. BRANNOCK

Telephone No. (703) 308-0196